

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 10:26:01 ; Search time 23.43 seconds
(without alignments)
178.687 Million cell updates/sec

Title: US-08-486-814-17
Perfect score: 1080
Sequence: 1 MSRLDKSVINSALELLNEV.....FGLELIICGLEKQKCESGS 207

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database : A_Geneseq_34.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1080	100.0	335	1	R64808	tTA transactivator
2	1080	100.0	207	1	R89763	Wild type Tn-10 de
3	1080	100.0	207	1	W08296	Wild-type E.coli T
4	1080	100.0	651	1	W48357	Multi-chimeric tra
5	1080	100.0	651	1	W47583	Multi-chimeric tra
6	1080	100.0	207	1	W1326	Protein encoded by
7	1075	99.5	207	1	R48630	Sequence of the te
8	1075	99.5	297	1	R64809	tTAS transactivato
9	1075	99.5	297	1	R85324	tTAS transactivato
10	1075	99.5	1088	1	R88636	Plasmid pASK75 ope
11	1075	99.5	354	1	W21994	Tetracycline trans
12	1075	99.5	349	1	W54311	Plectstrin homolog
13	1073	99.4	335	1	R53323	tTA transactivator
14	1070	99.1	336	1	W08474	pret-tTAK encoded
15	1068	98.9	207	1	W08323	Mutated Tet repres
16	1067	98.8	207	1	W08317	Mutated Tet repres
17	1067	98.8	207	1	W08318	Mutated Tet repres
18	1067	98.8	207	1	W08300	Mutated Tet repres
19	1066	98.7	207	1	W08320	Mutated Tet repres
20	1066	98.7	207	1	W08304	Mutated Tet repres
21	1065	98.6	207	1	W08312	Mutated Tet repres
22	1063	98.4	207	1	W08306	Mutated Tet repres
23	1061	98.2	207	1	W08325	Mutated Tet repres
24	1061	98.2	207	1	W08326	Mutated Tet repres
25	1059	98.1	207	1	W08301	Mutated Tet repres
26	1058	98.0	335	1	R89761	Mutated Tn-10 deri
27	1058	98.0	207	1	R89764	Mutated Tn-10 deri
28	1058	98.0	335	1	W08293	Mutated reverse Te
29	1058	98.0	207	1	W08297	Mutated Tet repres
30	1058	98.0	207	1	W08321	Mutated Tet repres
31	1058	98.0	207	1	W08305	Mutated Tet repres
32	1058	98.0	335	1	W71323	Protein encoded by
33	1058	98.0	207	1	W71327	Protein encoded by
34	1057	97.9	207	1	W08303	Mutated Tet repres
35	1057	97.9	207	1	W08310	Mutated Tet repres
36	1057	97.9	207	1	W08313	Mutated Tet repres
37	1056	97.8	207	1	W08316	Mutated Tet repres
38	1056	97.8	207	1	W08319	Mutated Tet repres
39	1056	97.8	207	1	W08302	Mutated Tet repres
40	1056	97.8	207	1	W08308	Mutated Tet repres
41	1055	97.7	207	1	W08322	Mutated Tet repres
42	1055	97.7	207	1	W08324	Mutated Tet repres
43	1055	97.7	207	1	W08307	Mutated Tet repres

ALIGNMENTS

RESULT 1

R64808
ID R64808 standard; Protein; 335 AA.
AC R64808:
DT 17-JUL-1995 (first entry)
DE tTA transactivator.
KW tTA; transactivator; tetracycline-controllable transactivator;
KW conditional inactivation; homologous recombination; gene expression;
KW gene regulation; gene therapy; tetracycline-resistance; tetr;
KW transgenic animal.
OS Herpes simplex virus K12, KOS.
PN W09429442-A.
PD 22-DEC-1994.
PF 14-JUN-1994; U06734.
PR 14-JUN-1993; US-076327.
PA (BADI) BASF AG.
PI Bujard H, Gossen M, Salfeld JG, Voss JW;
DR WPI: 95-036472/05.
DR N-PSDB; Q76264.
PT Regulatory systems using tetracycline-controllable transactivator
PT (tTA) - useful for conditional inactivation or modulation of
PT gene expression in a host cell or animal.
PS Disclosure: Page 46-48; 103pp; English.
CC A 397 bp Mu1/Pok1 fragment of pMSVP16 coding for the C-terminal 130
CC amino acids of VP16 of HSV was blunted and inserted in pUHD14-1.
CC The resulting plasmid, pUHD15-1, encoded a tetr-VP16 fusion protein,
CC or tetracycline-controllable transactivator (tTA), whose sequence is
CC given in Q76264, and the encoded protein sequence in R64808.
SQ Sequence 335 AA;

Query Match 100.0%; Score 1080; DB 1; Length 335;

Best Local Similarity 100.0%; Pred. No. 1.2e-112; Indels 0; Gaps 0;
Matches 207; Conservative 0; Mismatches 0;

QY 1 MSRLDKSVINSALELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
Db 1 MSRLDKSVINSALELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
QY 61 DRHHTFCPLEGESWQDFLRNKAQKSPFCALLSHRDGAKVHLGTRPTEKQVETLENOIAFL 120
Db 61 DRHHTFCPLEGESWQDFLRNKAQKSPFCALLSHRDGAKVHLGTRPTEKQVETLENOIAFL 120
QY 121 CQQGFLENALYALSAYGHFTLGCVLDDQEHQVAKEREPTTDSMPPLLRQAIELFDHQ 180
Db 121 CQQGFLENALYALSAYGHFTLGCVLDDQEHQVAKEREPTTDSMPPLLRQAIELFDHQ 180
QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207
Db 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 2

R89763
ID R89763 standard; Protein; 207 AA.
AC R89763:
DT 07-JUL-1996 (first entry)
DE Wild type Tn-10 derived Tet repressor.
KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
KW fusion protein; gene expression; regulation; inhibition; activation;
KW transcription.
OS Transposon Tn-10.
PN W09601313-A1.
PD 18-JAN-1996.
PR 29-JUN-1995; U08179.
PR 01-JUL-1994; US-270637.

PR 15-JUL-1994; US-275876.
 PR 03-FEB-1995; US-383754.
 PR 07-JUN-1995; US-486814.
 PA (BUJA/) BUJARD H.
 PA (GOSS/) GOSSSEN M.
 PI Bujard H, Gossen M;
 DR WPI; 96-087666/09.
 DR N-PSDB; T11351.
 PR New tetracycline-regulated transcription modulators - comprising
 PT fusion proteins which bind to tet operator sequences to activate or
 PT inhibit transcription
 PS Claim 38; Page 78; 112pp; English.
 CC Fusion proteins comprising a first polypeptide which binds to a tet
 CC operator sequence in the presence of tetracycline or a tetracycline
 CC analogue, operatively linked to a second polypeptide which either
 CC activates or inhibits transcription in eukaryotic cells may be used
 CC to activate or inhibit transcription. Such proteins may be used to
 CC regulate gene expression in cells and may be particularly useful for
 CC gene therapy and for expression of gene products in transgenic
 CC organisms. Induction of gene expression is rapid, efficient and
 CC strong, typically 1000-2000 fold. The inducing agent does not cause
 CC pleiotropic effects or cytotoxicity in eukaryotic cells. This
 CC sequence is the wild type Tn-10 tet repressor.
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred No. 5.9e-113;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 DB 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALDLAIEML 120
 DB 61 DRHHTFCPLEGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALDLAIEML 120
 QY 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREPTPTDSMPLLRQAIELFDHQ 180
 DB 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREPTPTDSMPLLRQAIELFDHQ 180
 QY 181 GAEPALFGLLELICGLKOLKCESGS 207
 DB 181 GAEPALFGLLELICGLKOLKCESGS 207

RESULT 3
 ID W08296 standard; Protein; 207 AA.
 AC W08296;
 DT 19-MAR-1997 (first entry)
 DE Wild-type E.coli Tn10-derived Tet repressor.
 KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;
 KW repressor; gene expression; therapy; transgenic animal; disease model;
 KW Tn10; transposon 10.
 OS Escherichia coli.
 PN W08640892-Al.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09049.
 PR 07-JUN-1995; US-485971.
 PA (BADI) BASF AG.
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;
 DR WPI; 97-052305/05.
 PT Nucleic acid encoding tetracycline-inducible transcription
 PT regulatory fusion protein - comprising modified tetracycline
 PT repressor able to bind mutant tet operator, fused to transcription
 PT regulator, useful for modulating eukaryotic gene expression
 PS Example 4; Page 84-85; 117pp; English.
 CC W08296 represents wild-type tet repressor (TetR) protein, derived from
 CC transposon 10 (Tn10) of E. coli. The wild-type sequence is used for
 CC the production of a mutated TetR (e.g. so as to bind its target in the
 CC presence rather than the absence of tetracycline). Mutant TetR proteins

CC can be fused to a transcriptional activator e.g. VP16 (herpes simplex
 CC virus virion protein 16), to form a Tc-controlled transcriptional
 CC activator (tTA). The main invention of the specification concerns
 CC modified TetR proteins that bind to modified class B tet operator
 CC sequences tetO-4C and tetO-6C (see T45711 and T48478). Modified TetR
 CC proteins can be fused to any transcription regulatory polypeptide and
 CC used to control transcription of a tetO-4C or tetO-6C linked gene.
 CC Nucleic acid encoding such a fusion protein may be introduced into a
 CC cell and transcription of the protein can be controlled by altering the
 CC concn of tetracycline (or an analogue) in the cell, as appropriate.
 CC This ability to modulate gene expression in a predictable way is very
 CC useful in gene therapy and for recombinant protein prodn. in cultured
 CC cells or transgenic animals. The Tc-inducible system is also useful for
 CC the prodn. of transgenic animal models for the study of disease and
 CC also for the study of gene function e.g. during differentiation. The
 CC Tc-inducible system allows rapid activation of gene transcription
 CC without cellular toxicity, high concns. of inducer are not required.
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred No. 5.9e-113;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 DB 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALDLAIEML 120
 DB 61 DRHHTFCPLEGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALDLAIEML 120
 QY 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREPTPTDSMPLLRQAIELFDHQ 180
 DB 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREPTPTDSMPLLRQAIELFDHQ 180
 QY 181 GAEPALFGLLELICGLKOLKCESGS 207
 DB 181 GAEPALFGLLELICGLKOLKCESGS 207

RESULT 4
 ID W48357 standard; Protein; 651 AA.
 AC W48357;
 DT 20-JUL-1998 (first entry)
 DE Multi-chimeric transactivating factor tTAER fusion protein.
 KW Multi-chimeric transactivating factor; tTAER; tetR;
 KW tetracycline repressor; HSV; oestrogen receptor; promoter;
 KW packaging cell line; retrovirus; retroviral particle; vector;
 KW gene delivery; gene therapy.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Herpes simplex virus.
 OS Chimeric - Mammalia.
 PN W09805759-Al.
 PD 12-FEB-1998.
 PF 06-AUG-1997; U13846.
 PR 07-AUG-1996; US-594652.
 PA (CITY) CITY OF HOPE.
 PA (REGC) UNIV CALIFORNIA.
 PI Chen S, Friedmann T, Miyanochara A, Prussak CE, Yee J;
 DR WPI; 98-145602/13.
 DR N-PSDB; V17756.
 PT New packaging cell lines for pseudotyped retroviral vectors -
 PT comprises sequences encoding retroviral gag and pol polypeptide(s)
 PT and envelope protein
 PS Disclosure; Fig 4A-C; 84pp; English.
 CC This fusion protein comprises a multi-chimeric transactivator,
 CC designated tTAER, that is composed of (from the N-terminus to the
 CC C-terminus) the Escherichia coli tetR polypeptide, the
 CC transcriptional activation domain of herpes simplex virus VP16, and
 CC the ligand binding domain of the oestrogen receptor. An inducible
 CC expression system of the invention is composed of tTAER and a

CC minimal promoter (see V17755) derived from the immediate early gene
CC of cytomegalovirus linked to 7 tandem copies of the tet operator
CC (teto) that is the binding site for tetr, which in turn can be
CC linked to a nucleotide sequence of interest. The invention relates
CC to packaging cell lines derived from HeLa, D17, MDCK, BHK or
CC preferably Cf2th cells and recombinant retroviral particles
CC produced by them, particularly pseudotyped retroviral particles.
CC Retroviral particles are produced by inducibly expressing an
CC envelope protein by linking an envelope protein-encoding nucleotide
CC sequence to the inducible expression system. The products can be
CC used for the inducible expression in cells of polypeptides, e.g.
CC cytotoxic products or therapeutic agents. The activation of the
CC inducible expression system requires 2 independent signals, thus
CC reducing the incidence of undesired transcriptional activation.
SQ Sequence 651 AA;

Query Match 100.0%; Score 1080; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 2.9e-112;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
Db 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60

QY 61 DRHHTFCPLGEGSWQDFLRNKAISFRALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
Db 61 DRHHTFCPLGEGSWQDFLRNKAISFRALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

QY 121 CQGFSLLENALYALS AVGHTFLGCVLEQDQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180
Db 121 CQGFSLLENALYALS AVGHTFLGCVLEQDQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180

QY 181 GAEPFLFGLLEIICGLEKQKCESGS 207
Db 181 GAEPFLFGLLEIICGLEKQKCESGS 207

RESULT 5

ID W47583 standard; Protein; 651 AA.

AC W47583; 1998 (first entry)

DT Multi-chimeric transactivator tTAER sequence.

DE Inducible expression system; modulation; cytotoxin; therapeutic;

KW tTAER; multi-chimeric; transactivator; tetR; Vp16;

KW ligand-binding domain.

OS Escherichia coli.

OS Herpes simplex virus.

PN W09805754-A2.

PD 12-FEB-1998.

PF 29-JUL-1997; U13221.

PR 07-AUG-1996; US-693940.

PA (CITY) CITY OF HOPE.

PA (REGC) UNIV CALIFORNIA.

PI Chen ST, Friedmann T, Yee JK;

DR WPI; 98-145597/13.

DR N-PSDB; V18690.

PT New inducible expression systems - comprising multi-chimeric
PT transactivator, induces transcription from promoter in the absence
PT of first ligand and presence of second ligand

PS Example 1; Fig 4; 72pp; English.

CC The sequence is that of multi-chimeric transactivator tTAER.
CC This is a fusion protein comprising, from N-terminus to C-terminus,
CC the E.coli tetr polypeptide, the transcription activation domain
CC of HSV Vp16 and the ligand-binding domain of an oestrogen receptor
CC from pHE14 (Kumar et al. 1986). DNA encoding it can be used as part of
CC an inducible expression system which induces transcription from the
CC promoter in the absence of the first ligand and presence of the
CC second. The products can be used for the inducible expression in cells
CC of polypeptides such as cytotoxic products or therapeutic products.
CC The activation of the inducible expression system requires 2
CC independent signals, which reduces the incidence of undesired

CC transcriptional activation.
SQ Sequence 651 AA;

Query Match 100.0%; Score 1080; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 2.9e-112;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
Db 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60

QY 61 DRHHTFCPLGEGSWQDFLRNKAISFRALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
Db 61 DRHHTFCPLGEGSWQDFLRNKAISFRALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

QY 121 CQGFSLLENALYALS AVGHTFLGCVLEQDQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180
Db 121 CQGFSLLENALYALS AVGHTFLGCVLEQDQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180

QY 181 GAEPFLFGLLEIICGLEKQKCESGS 207
Db 181 GAEPFLFGLLEIICGLEKQKCESGS 207

RESULT 6

ID W71326 standard; Protein; 207 AA.

AC W71326;

DT 04-DEC-1998 (first entry)

DE Protein encoded by wild type Tn10-derived tet repressor.

KW Tet repressor; tetracycline; regulation; expression;

KW Tet operator-linked gene; tet operator.

OS Unidentified.

PN US5814618-A.

PD 29-SEP-1998.

PF 07-JUN-1995; 485978.

PR 07-JUN-1995; US-485978.

PR 14-JUN-1993; US-076327.

PR 14-JUN-1993; US-076726.

PR 14-JUN-1994; US-260452.

PR 01-JUL-1994; US-270637.

PR 15-JUL-1994; US-275876.

PR 06-FEB-1995; US-383754.

PA (BADI) BASF AG.

PA (KNOL) KNOLL AG.

PI Bujard H, Gossen M;

DR WPI; 98-541795/46.

DR N-PSDB; V60088.

PT Tetracycline based regulation of gene expression - uses a
PT tetracycline operator sequence joined to a gene of interest, the
PT gene of interest being induced in the presence, but not absence of
PT the antibiotic

PS Claim 4; Columns 77-80; 63pp; English.

CC The present sequence is encoded by wild type Tn10-derived tet repressor.
CC It is used in the course of the invention. The specification describes a
CC method for regulating expression of a Tet (tetracycline) operator-linked
CC gene in a cell of a subject. The method comprises introducing into the
CC cell a nucleic acid encoding a fusion protein which inhibits
CC transcription in eukaryotic cells, the fusion protein comprising a
CC polypeptide which binds to a Tet operator sequence, operatively linked
CC to heterologous second polypeptide which inhibits transcription in
CC eukaryotic cells and modulating the concentration of a tetracycline
CC (analogue) in the subject. The method is used for the regulation of
CC gene expression system, using tetracycline (analogues). The system
CC enables a gene coupled to the system to be induced in the presence of
CC Tet and then stopped when Tet is removed.

SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 5.9e-113;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALATEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALATEML 60
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 QY 121 COQGSLENALYALS AVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
 DB 121 COQGSLENALYALS AVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207
 DB 181 GAEPALFGLLELIICGLEKQKCESGS 207

RESULT 7
 R48630
 ID R48630 standard; Protein; 207 AA.
 AC R48630;
 DT 19-AUG-1994 (first entry)
 DE Sequence of the tetracycline repressor tetr.
 KW Tetracycline repressor; TetR; Tn10; transposon 10.
 OS Escherichia coli.
 PN W09404672-A.
 PD 03-MAR-1994.
 PF 26-AUG-1993; U08230.
 PR 26-AUG-1992; US-935763.
 PA (DNXXD-) DNK CORP.
 PI Byrne G;
 DR WPI; 94-083191/10.
 DR N-PSDB; Q56710.
 FT Tetracycline repressor-mediated regulation system - useful for
 PT controlling gene expression in transgenic animals
 PS Disclosure; Page 41-42; 76pp; English.
 CC The inventors claim a construct which comprises an animal promoter
 CC element having a tetracycline repressor (tetR) operator. The
 CC promoter element may be the PEPC promoter which is tissue specific
 CC being expressed selectively in the liver and becoming active
 CC shortly prior to birth. The tetR sequence in the
 CC construct is 3' to a TATA-box sequence and is inserted into the NheI
 CC site of the PEPC promoter element. The entire sequence of the tetR
 CC repressor is given in Q56710/R48630.
 SQ Sequence 207 AA;

Query Match 99.5%; Score 1075; DB 1; Length 207;
 Best Local Similarity 99.5%; Pred. No. 2.1e-112;
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALATEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALATEML 60
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 QY 121 COQGSLENALYALS AVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
 DB 121 COQGSLENALYALS AVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207
 DB 181 GAEPALFGLLELIICGLEKQKCESGS 207

RESULT 8
 R64809
 ID R64809 standard; Protein; 297 AA.
 AC R64809;

DT 17-JUL-1995 (first entry)
 DE tTAs transactivator.
 KW tTAs; transactivator; tetracycline-controllable transactivator;
 KW conditional inactivation; homologous recombination; gene expression;
 KW gene regulation; gene therapy; tetracycline-resistance; tetR;
 KW transgenic animal.
 OS Herpes simplex virus K12, KOS.
 PN W09429442-A.
 PD 22-DEC-1994.
 PF 14-JUN-1994; U06734.
 PR 14-JUN-1993; US-076327.
 PA (BADI) BASF AG.
 PI Bujard H, Gossen M, Salfeld JG, Voss JW;
 DR WPI; 95-036472/05.
 DR N-PSDB; Q76265.
 FT Regulatory systems using tetracycline-controllable transactivator
 PT (tTA) - useful for conditional inactivation or modulation of
 PT gene expression in a host cell or animal
 PS Disclosure; Page 50-51; 103pp; English.
 CC A DNA fragment of pMSVP16 coding for the C-terminal 97 amino
 CC acids of VP16 of HSV was blunted and inserted in pUHD14-1. The
 CC resulting plasmid, pUHD151-1, encoded a tetR-VP16 fusion protein,
 CC or tetracycline-controllable transactivator (smaller version, tTAs),
 CC whose sequence is given in Q76265.
 SQ Sequence 297 AA;

Query Match 99.5%; Score 1075; DB 1; Length 297;
 Best Local Similarity 99.5%; Pred. No. 3.5e-112;
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALATEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALATEML 60
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 QY 121 COQGSLENALYALS AVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
 DB 121 COQGSLENALYALS AVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207
 DB 181 GAEPALFGLLELIICGLEKQKCESGS 207

RESULT 9
 R85324
 ID R85324 standard; Protein; 297 AA.
 AC R85324;
 DT 13-MAR-1996 (first entry)
 DE tTAs transactivator.
 KW Transactivator; tTAs; tet repressor; tetR; virion protein 16; VP16;
 KW gene expression; tetracycline-responsive promoter;
 KW transgenic animal
 OS Chimeric Escherichia coli;
 OS Chimeric Herpes simplex virus.
 PN US5464758-A.
 PD 07-NOV-1995.
 PR 14-JUN-1993; 076726.
 PR 14-JUN-1993; US-076726.
 PA (BUJA/) BUJARD H.
 PA (GOSS/) GOSSEN M.
 PI Bujard H, Gossen M;
 DR WPI; 95-392612/50.
 DR N-PSDB; T06868.
 PT Polynucleotide encoding transactivator fusion protein contg. tet
 PT repressor - used to control expression of gene regulated by minimal
 PT promoter linked to tet operon, and vectors and cells where gene
 PT expression is regulated by tetracycline
 PS Disclosure; Fig 5A-B; 37pp; English.

CC A fusion protein (R85324) composed of the Escherichia coli tn10
 CC tet repressor (tetR) and a 127-amino acid C-terminal portion of the
 CC transcription activating domain of herpes simplex virus virion
 CC protein 16 (VP16) is the product of novel tetracycline-controlled
 CC transactivator tTA (T06868). The transactivator is used to regulate
 CC expression of a heterologous gene operably linked to a minimal promoter
 CC derived from human cytomegalovirus promoter IE (see T06869-70) and at
 CC least 1 tet operator (tetO) sequence. On/off regulation of expression
 CC of the heterologous protein by host eukaryotic cells is provided by
 CC varying the medium tetracycline conc. Transgenic animals producing
 CC a desired protein in their milk can also be produced.
 CC Sequence 297 AA;

Query Match 99.5%; Score 1075; DB 1; Length 297;
 Best Local Similarity 99.5%; Pred. NO. 3.5e-112;
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60
 DB 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 QY 121 CQOGFSLENALYALSVAHGFTLCVLEDOHQVAKEREETPTTDSMPPLLROAIELFDHQ 180
 DB 121 CQOGFSLENALYALSVAHGFTLCVLEDOHQVAKEREETPTTDSMPPLLROAIELFDHQ 180
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207
 DB 181 GAEPALFGLLELIICGLEKQKCESGS 207

RESULT 10
 R88636
 ID R88636 standard; Protein; 1088 AA.
 AC R88636;
 DT 27-JUN-1996 (first entry)
 DE Plasmid pASK75 open reading frame (c) translation.
 KW Tetracycline; resistance; TetR gene; repressor protein; TetA gene;
 KW prokaryotic expression vector; beta-lactamase promoter; bla gene;
 KW circular; cyclic.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_difference 1..1088 /note="each X corresponds to a termination codon"
 FT peptide 49..69 /label="OmpA_leader"
 FT protein 614..820 /label="Tetr"
 FT region 601..609 /note="tetracycline repressor"
 FT /note="C-terminal end of beta-lactamase; the
 FT exact position of the N-terminus is not
 FT given in the specification"
 FT W09532295-AL.
 PN 30-NOV-1995.
 PE 17-MAY-1995; E01862.
 PR 19-MAY-1994; DE-417598.
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PI Skerra A, Wardenberg C;
 DR N-PSDB; T11192.
 DR Prokaryotic vector for regulated prodn. of heterologous protein gene
 PT - controlled by promoter repressed by tetracycline repressor
 PT protein, prodn. of which is controlled by independent promoter not
 PT regulated by this repressor
 PS Disclosure; Fig 1a; 50pp; German.
 CC Plasmid pASK75 is a specifically claimed example of a prokaryotic
 CC expression vector comprising the tetracycline promoter/operator (P/O)
 CC region and the tetracycline repressor (tetR) gene. The tetR gene is

CC under the control of the beta-lactamase (bla) promoter. Expression of
 CC heterologous genes inserted downstream of the tet P/O is controlled
 CC by the activity of the Tet repressor. Plasmid pASK75 was derived
 CC from pASK60-Strep (Schmidt and Skerra, Protein. Eng. 6 (1993), 109-
 CC 122) and further contains the OmpA leader peptide coding sequence, a
 CC polylinker and a streptavidin-binding peptide strep-tag coding
 CC region from the parent plasmid. In addition, pASK75 contains the
 CC ColEI origin of replication, the bla gene, intergene regions of
 CC filamentous phage f1 and a lipoprotein terminator.
 CC The specification includes the translated sequences from all three
 CC reading frames of pASK75; the OmpA leader, TetR and Bla coding
 CC regions are all in reading frame (c).
 CC Sequence 1088 AA;

Query Match 99.5%; Score 1075; DB 1; Length 1088;
 Best Local Similarity 99.5%; Pred. NO. 2.2e-111;
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60
 DB 614 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 673
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 674 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 733
 QY 121 CQOGFSLENALYALSVAHGFTLCVLEDOHQVAKEREETPTTDSMPPLLROAIELFDHQ 180
 DB 734 CQOGFSLENALYALSVAHGFTLCVLEDOHQVAKEREETPTTDSMPPLLROAIELFDHQ 793
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207
 DB 794 GAEPALFGLLELIICGLEKQKCESGS 820

RESULT 11
 W21994
 ID W21994 standard; Protein; 354 AA.
 AC W21994;
 DT 26-JAN-1998 (first entry)
 DE Tetracycline transactivator fusion protein tetRIELA.
 KW Baculovirus; recombinant protein; neurotoxin; parasite;
 KW biological control; insecticide; insecticidal protein;
 KW nuclear polyhedrosis virus; early promoter; tetracycline.
 OS Chimeric - Baculovirus.
 OS Chimeric - Escherichia coli.
 FH Key Location/Qualifiers
 FT Peptide 1..209 /label="tetr peptide"
 FT Peptide 210..354 /label="IEIA peptide"
 FT W09723636-AL.
 PN 03-JUL-1997.
 PE 16-DEC-1996; U20512.
 PR 12-DEC-1996; US-764369.
 PR 22-DEC-1995; US-009120.
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 PI McCutchen BF;
 DR WPI; 97-351071/32.
 DR Controlling expression of insecticidal proteins - using recombinant
 PT baculovirus containing tetracycline transactivator protein
 PS Example 2; Fig 8; 63pp; English.
 CC This transactivator protein is encoded by a hybrid gene consisting of
 CC the tetracycline repressor gene (tetR) of E. coli (T74611) fused
 CC in frame with the immediate early promoter (IEIA) of baculovirus
 CC (T74614). The protein is used to demonstrate a novel method of gene
 CC expression regulation of recombinant baculoviruses in an insect
 CC cell/host. An insect cell is constructed having a chimeric gene, which
 CC consists of a promoter (e.g. T74614) operably linked to DNA encoding a
 CC regulatory protein capable of affecting gene expression. The regulatory
 CC protein is preferably a tetracycline transactivator protein
 CC (e.g. W21994). A recombinant baculovirus expression vector is then

CC constructed. This has a second chimeric gene, which consists of a
 CC promoter (one or more tetracycline operator sites operably linked to a
 CC minimal promoter region, see T74615-16) that is affected by the above
 CC regulatory protein. This promoter is operably linked to DNA encoding an
 CC insecticidal protein, (w21993). The recombinant baculovirus expression
 CC vector is introduced into the recombinant insect cell. Expression of the
 CC regulatory protein affects expression of the insecticidal protein. The
 CC recombinant insect cells and the recombinant baculoviruses may be used
 CC in the production of insecticidal recombinant baculoviruses. Insect cells
 CC treated with recombinant viruses, as described, show an increase in the
 CC yield of polyhedral inclusion bodies.
 SQ Sequence 354 AA;

Query Match 99.5%; Score 1075; DB 1; Length 354;
 Best Local Similarity 99.5%; Pred. No. 4.5e-112;
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 60
 DB 1 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 60
 QY 61 DRHHTHFCPLGESQWDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 61 DRHHTHFCPLGESQWDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 QY 121 COQGSLENALYALSAGVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180
 DB 121 COQGSLENALYALSAGVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180
 QY 181 GAEPALFGLLELIICGLEKQKLCESGS 207
 DB 181 GAEPALFGLLELIICGLEKQKLCESGS 207

RESULT 12
 W54311
 ID W54311 standard; Protein; 349 AA.
 AC W54311;
 DE 28-AUG-1998 (first entry)
 DT Pleckstrin homology domain-tet repressor fusion.
 KW Pleckstrin homology domain; PH domain; peptide library;
 KW functional molecular surface; protein structural template;
 KW vaccine; gene therapy; cytohesin 1; human; plasmid pTLP2;
 KW tet repressor.
 OS Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Synthetic.
 PH Key Location/Qualifiers
 FT 1..221
 FT /label= Tet_repressor
 FT 222..349
 FT /label= PH_domain
 PN W09745538-A1.
 PD 04-DEC-1997.
 PF 30-MAY-1997; E02840.
 PR 31-MAY-1996; EP-108776.
 PA (MEDI-) MEDIGENE AG.
 PI Bruhn H, Funk M, Henkel T, Steipe B;
 DR WPI: 98-230215/20.
 DR N-PSDB: V26524.
 PT Vectors used to produce PH domain-like peptide libraries - which are
 PT screened for therapeutically useful peptide(s), e.g. to produce
 PT vaccines
 PS Example 4; Page 79-81; 137pp; English.
 CC This polypeptide comprises a fusion between a synthetic Pleckstrin
 CC homology (PH) domain (see W54285) of human cytohesin 1 and a tet
 CC repressor. It is encoded by plasmid pTLP2 (see V26524). The fusion
 CC protein is expressed from a polycistronic message, constitutively
 CC transcribed from the beta-lactamase promoter. The plasmid was used
 CC to transform Escherichia coli JM109 cells. The construct was used
 CC to demonstrate the folding, stability, functionality and ease of
 CC handling of a fusion protein comprising a synthetic PH domain and

CC a DNA binding domain. The invention provides vectors that are used
 CC for the production of PH domain-like peptide libraries, which can
 CC be screened to identify peptides that have desirable properties,
 CC especially novel binding or catalytic properties, and which may be
 CC of use in research or therapy, or as vaccines. Novel synthetic
 CC protein structural templates for the generation, screening and
 CC evolution of functional molecular surfaces are provided.
 SQ Sequence 349 AA;

Query Match 99.5%; Score 1075; DB 1; Length 349;
 Best Local Similarity 99.5%; Pred. No. 4.4e-112;
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 60
 DB 2 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 61
 QY 61 DRHHTHFCPLGESQWDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 62 DRHHTHFCPLGESQWDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 121
 QY 121 COQGSLENALYALSAGVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180
 DB 122 COQGSLENALYALSAGVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 181
 QY 181 GAEPALFGLLELIICGLEKQKLCESGS 207
 DB 182 GAEPALFGLLELIICGLEKQKLCESGS 208

RESULT 13
 R85323
 ID R85323 standard; Protein; 335 AA.
 AC R85323;
 DT 13-MAR-1996 (first entry)
 DE TTA transactivator.
 KW Transactivator; tTA; tet repressor; tetR; virion protein 16; VP16;
 KW gene expression; tetracycline-responsive promoter;
 KW transgenic animal.
 OS Chimeric Escherichia coli;
 OS Chimeric Herpes simplex virus.
 PN US5464758-A.
 PD 07-NOV-1995.
 PF 14-JUN-1993; 076726.
 PR 14-JUN-1993; US-076726.
 PA (BUJA/) BUJARD H.
 PA (GOSSE/) GOSSEN M.
 PI Bujard H, Gossen M;
 DR WPI: 95-392612/50.
 DR N-PSDB: T06867.
 PT Polynucleotide encoding trans:activator fusion protein contg. tet
 PT repressor - used to control expression of gene regulated by minimal
 PT promoter linked to tet operon, and vectors and cells where gene
 PT expression is regulated by tetracycline
 PS Disclosure; Fig 4A-B; 37pp; English.
 CC A fusion protein (R85323) composed of the Escherichia coli Tn10
 CC tet repressor (tetR) and a 127-amino acid C-terminal portion of the
 CC transcription activating domain of herpes simplex virus virion
 CC protein 16 (VP16) is the product of novel tetracycline-controlled
 CC transactivator tTA (T06867). The transactivator is used to regulate
 CC expression of a heterologous gene operably linked to a minimal promoter
 CC derived from human cytomegalovirus promoter IE (see T06869-70) and at
 CC least 1 tet operator (tetO) sequence. On/off regulation of expression
 CC of the heterologous protein by host eukaryotic cells is provided by
 CC varying the medium tetracycline conc. Transgenic animals producing
 CC a desired protein in their milk can also be produced.
 SQ Sequence 335 AA;

Query Match 99.4%; Score 1073; DB 1; Length 335;
 Best Local Similarity 99.5%; Pred. No. 7e-112;
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 DB 1 MSRLDKSVINSALLELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 QY 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 QY 121 CQOQFSLENALYALSAGVHFTLGCVLDEQHQVAKEREPTTDSMPPLRQALIEFDHQ 180
 DB 121 CQOQFSLENALYALSAGVHFTLGCVLDEQHQVAKEREPTTDSMPPLRQALIEFDHQ 180
 QY 181 GAEPALFGLLEIICGLEKQKCESSG 207
 DB 181 GAEPALFGLLEIICGLEKQKCESSG 207

RESULT 14

W08474
 ID W08474 standard; Protein; 336 AA.
 AC W08474;
 DT 28-MAR-1997 (first entry)
 DE pTet-trak encoded protein.
 KW pTet-trak; pTet-Splice; tetracycline transactivator; promoter;
 KW gene expression; transgenic animal; animal model; drug screening.
 OS Synthetic.
 PN W09640946-Al.
 PD 19-DEC-1996.
 PF 07-JUN-1996; U10109.
 PR 07-JUN-1995; US-474169.
 PA (UYVA) UNIV YALE.
 PI Schatz DG;
 DR WPI; 97-077273/07.
 DR N-PSDB; T49877.
 DR Nucleic acid encoding tetracycline transactivator fusion protein -
 PT provides rapid and reversible control of gene expression, e.g. for
 PT creating animal models for drug screening
 PS Example 1; Fig 10b-g; 82pp; English.
 CC Autoregulatory plasmid pTet-trak (T49877) contains a modified
 CC tetracycline transactivator gene, trak, under control of a minimal
 CC promoter. It encodes a 366-amino acid protein (W08474). An
 CC autoregulatory tetracycline-regulated system for inducible gene
 CC expression in eukaryotes has been developed that allows the rapid,
 CC reversible control of gene expression in eukaryotic cells, and
 CC creation of transgenic animals in which expression of a luciferase
 CC reporter gene is controlled by altering the concentration of
 CC tetracycline in the drinking water of the animals.
 SQ Sequence 336 AA;

Query Match 99.1%; Score 1070; DB 1; Length 336;
 Best Local Similarity 99.5%; Pred. No. 1.5e-111;
 Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SRDLKSVINSALLELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 61
 DB 3 SRDLKSVINSALLELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 62
 QY 62 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 121
 DB 63 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 122
 QY 122 CQOQFSLENALYALSAGVHFTLGCVLDEQHQVAKEREPTTDSMPPLRQALIEFDHQ 181
 DB 123 CQOQFSLENALYALSAGVHFTLGCVLDEQHQVAKEREPTTDSMPPLRQALIEFDHQ 182
 QY 182 GAEPALFGLLEIICGLEKQKCESSG 207
 DB 183 GAEPALFGLLEIICGLEKQKCESSG 208

RESULT 15

W08323
 ID W08323 standard; Protein; 207 AA.
 AC W08323;
 DT 21-MAR-1997 (first entry)
 DE Mutated Tet repressor for Tet-inducible transcriptional regulator.
 KW Tet; tetracycline; rc; operator; transcription; regulation; inducible;
 KW repressor; gene expression; therapy; transgenic animal; disease model.
 OS Synthetic.
 FH Key
 FT misc_difference 37 Location/Qualifiers
 FT /label= substitution
 FT /note= "Lys replaces wild-type Glu"
 FT misc_difference 39
 FT /label= substitution
 FT /note= "Gln replaces wild-type Pro"
 PN W09640892-Al.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09049.
 PR 07-JUN-1995; US-485971.
 PA (BADI) BASF AG.
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;
 DR WPI; 97-052305/05.
 DR Nucleic acid encoding tetracycline-inducible transcription
 PT regulatory fusion protein - comprising modified tetracycline
 PT repressor able to bind mutant tet operator, fused to transcription
 PT regulator, useful for modulating eukaryotic gene expression
 PS Claim 5; Page -; 117pp; English.
 CC W08300-W08326 are mutated Tet repressor (TetR) proteins. They were
 CC used to create fusion protein, comprising a mutated tetr fused to a
 CC transcriptional control sequence (e.g. v-erbA silencer domain), to
 CC produce tetracycline (Tc)-controlled transcriptional activators (ETA).
 CC The main invention of the specification concerns modified TetR proteins
 CC that bind to modified class B tet operator sequences tetO-4C and
 CC tetO-6C (see T45711 and T48478). Modified TetR proteins can be fused
 CC to any transcription regulatory polypeptide and used to control
 CC transcription of a tetO-4C or tetO-6C linked gene. Nucleic acid
 CC encoding such a fusion protein may be introduced into a cell and
 CC transcription of the protein can be controlled by altering the concn.
 CC of tetracycline (or an analogue) in the cell, as appropriate. This
 CC ability to modulate gene expression in a predictable way is very useful
 CC in gene therapy and for recombinant protein prodn. in cultured cells or
 CC transgenic animals. The Tc-inducible system is also useful for the
 CC prodn. of transgenic animal models for the study of disease and also
 CC for the study of gene function e.g. during differentiation. The
 CC Tc-inducible system allows rapid activation of gene transcription
 CC without cellular toxicity, high concns. of inducer are not required.
 SQ Sequence 207 AA;

Query Match 98.9%; Score 1068; DB 1; Length 207;
 Best Local Similarity 99.0%; Pred. No. 1.3e-111;
 Matches 205; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 DB 1 MSRLDKSVINSALLELLNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 QY 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120
 QY 121 CQOQFSLENALYALSAGVHFTLGCVLDEQHQVAKEREPTTDSMPPLRQALIEFDHQ 180
 DB 121 CQOQFSLENALYALSAGVHFTLGCVLDEQHQVAKEREPTTDSMPPLRQALIEFDHQ 180
 QY 181 GAEPALFGLLEIICGLEKQKCESSG 207
 DB 181 GAEPALFGLLEIICGLEKQKCESSG 207

Search completed: June 9, 1999, 12:59:17
 Job time: 9196 sec

Fri Jun 11 11:58:05 1999

us-08-486-814-17.rag

Page 8

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 10:27:01 ; Search time 19.42 Seconds
(without alignments)
399.294 Million cell updates/sec

Title: US-08-486-814-17
Perfect score: 1080
Sequence: 1 MSRLDKSVINSALLELNEV.....FGLELIICGLEKQLKCESGS 207
Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database : PIR_58:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1075	99.5	207	1	RPBCTN	repressor tetr - E
2	683	63.2	218	2	S07359	regulatory protein
3	683	63.2	218	2	S30287	regulatory protein
4	619	57.3	126	2	A26948	tetracycline resis
5	525	48.6	219	1	RPECY5	tetracycline repre
6	515.5	47.7	216	2	JQ1478	regulatory protein
7	515.5	47.7	216	2	S38655	tetr protein - Pse
8	509.5	47.2	211	2	I77569	tet repressor - Es
9	505.5	46.8	216	1	RPECR1	tet repressor prot
10	176.5	16.3	261	2	H70740	hypothetical prote
11	149.5	13.8	259	2	A40046	Tetr repressor hom
12	119	11.0	191	2	A69900	hypothetical prote
13	110.5	10.2	210	2	S42417	probable transcrip
14	110.5	10.2	210	2	S42419	probable transcrip
15	108.5	10.0	189	2	H70042	transcription regu
16	101.5	9.4	192	2	C70487	transcription regu
17	98	9.1	196	2	B70827	hypothetical prote
18	95.5	8.8	196	2	D64918	glucuronide repres
19	94.5	8.8	236	2	S75298	hypothetical prote
20	94	8.7	179	2	B70391	transcription regu
21	93.5	8.7	246	2	E70861	hypothetical prote
22	92	8.5	213	2	F70946	probable regulato
23	89	8.2	194	2	C70035	transcription regu
24	87.5	8.1	235	2	A64695	conserved hypothet
25	86.5	8.0	192	2	E69280	iron-dependent rep
26	86	8.0	205	2	S29308	hypothetical prote
27	85.5	7.9	236	2	D64855	hypothetical prote
28	85.5	7.9	677	2	I49045	SKD3 - mouse
29	85	7.9	198	2	E69779	transcription regu
30	84	7.8	188	2	S12393	hypothetical prote
31	84	7.8	207	2	H70001	hypothetical prote
32	83.5	7.7	216	2	C70649	hypothetical prote
33	83	7.7	519	2	S31136	nitrogen fixation
34	83	7.7	195	2	S10899	regulatory protein
35	83	7.7	365	2	S17885	Tcd37 protein - fr
36	83	7.7	192	2	F70487	conserved hypothet
37	82	7.6	260	2	F64995	hypothetical prote
38	82	7.6	629	2	A71023	probable arginyl-t
39	81.5	7.5	299	2	S39744	transcription regu

RESULT 1
RPECTN
repressor tetr - Escherichia coli transposon Tn10
C:Species: Escherichia coli
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 12-Jun-1998
C:Accession: A03576; S02667
R:Postle, K.; Nguyen, T.T.; Bertrand, K.P.
Nucleic Acids Res. 12, 4849-4863, 1984
A:Title: Nucleotide sequence of the repressor gene of the Tn10 tetracycline resistance
A:Reference number: A03576; MUID:84247342
A:Accession: A03576
A:Molecule type: DNA
A:Residues: 1-207 <POS>
A:Cross-references: GB:X00694; NID:943051; PID:943052
R:Altschmid, L.; Baumeister, R.; Pfeleiderer, K.; Hillen, W.
EMBO J. 7, 4011-4017, 1988
A:Title: A threonine to alanine exchange at position 40 of tet repressor alters the r
A:Reference number: S02667; MUID:89091153
A:Accession: S02667
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-100 <ALT>
A:Cross-references: GB:J01830; NID:gl54845
C:Comment: This protein contains sequences similar to the DNA recognition regions of
C:Genetics:
A:Gene: tetr
C:Superfamily: tetracycline repressor
C:Keywords: antibiotic resistance; DNA binding; transcription regulation
F:25-46/Region: helix-turn-helix #status predicted

Query Match 99.5%; Score 1075; DB 1; Length 207;
Best Local Similarity 99.5%; Pred. No. 2.le-92;
Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
Db 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
QY 61 DRHHTHFCPLEGESWDQLRNKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQALFL 120
Db 61 DRHHTHFCPLEGESWDQLRNKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQALFL 120
QY 121 CQOGFSLENALYALSAVGHFTLGCVCLEQEHQVAKEREETPTTDSMPPLLRQALFQHQ 180
Db 121 CQOGFSLENALYALSAVGHFTLGCVCLEQEHQVAKEREETPTTDSMPPLLRQALFQHQ 180
QY 181 GAEPAPFLFGLLEIICGLEKQLKCESGS 207
Db 181 GAEPAPFLFGLLEIICGLEKQLKCESGS 207

RESULT 2
S07359
regulatory protein tetr - Escherichia coli plasmid RAI
N:Alternate names: tet repressor class D
C:Species: Escherichia coli
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 12-Jun-1998
R:Unger, B.; Klock, G.; Hillen, W.
Nucleic Acids Res. 12, 7693-7703, 1984
A:Title: Nucleotide sequence of the repressor gene of the RAI tetracycline resistance

A:Reference number: S07359; MUID:85037938

A:Accession: S07359

A:Molecule type: DNA

A:Residues: 1-218 <UNG>

A:Cross-references: EMBL:X01083; NID:g43053; PID:g43054

C:Genetics:

A:Gene: tetr

A:Genome: plasmid

C:Superfamily: tetracycline repressor

C:Keywords: antibiotic resistance; DNA binding; transcription regulation

F:25-46/Region: helix-turn-helix #status predicted

Query Match 63.2%; Score 683; DB 2; Length 218;
Best Local Similarity 65.2%; Pred. No. 4.1e-56;
Matches 131; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML 60

Db 1 MARLNRESVIDAALLELLNETGIDGLTRKLAQKLGIEQPTLYWHVKNKRALDLALAEVL 60

QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

Db 61 ARHHDYSLPAAGESWQDFLNKAKSFRALLRYRDGAKVHLGTRPDEKQYDVTETOLREM 120

QY 121 CQOGFSLNALYALSAGVHFTGLGCVLEDOEHQVAKKERETPTTDSMPPLLRQAIEFLDHQ 180

Db 121 TENGFSLRDLGLYALSASVSHFTLGAVLEQOEHTAALDRPAAPDENLPLLRQAIEFLDHQ 180

QY 181 GAEPAPFLGLELICGLEKOL 201

Db 181 DGEQAFHLGLESIRGFEVOL 201

RESULT 3

S30287

N:regulatory protein tetr - Salmonella ordonez plasmid pIP173

C:Species: Salmonella ordonez

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Jun-1998

C:Accession: S30287

R:Allard, J.D.; Gibson, M.L.; Vu, L.H.; Nguyen, T.T.; Bertrand, K.P.

Mol. Gen. Genet. 237, 301-305, 1993

A:Title: Nucleotide sequence of class D tetracycline resistance genes from Salmonella ordonez

A:Reference number: S30286

A:Accession: S30287

A:Molecule type: DNA

A:Residues: 1-218 <ALL>

A:Cross-references: EMBL:X65876; NID:g49073; PID:g49075

C:Genetics:

A:Gene: tetr

A:Genome: plasmid

C:Superfamily: tetracycline repressor

C:Keywords: antibiotic resistance; DNA binding; transcription regulation

F:25-46/Region: helix-turn-helix #status predicted

Query Match 63.2%; Score 683; DB 2; Length 218;
Best Local Similarity 65.2%; Pred. No. 4.1e-56;
Matches 131; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML 60

Db 1 MARLNRESVIDAALLELLNETGIDGLTRKLAQKLGIEQPTLYWHVKNKRALDLALAEVL 60

QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

Db 61 ARHHDYSLPAAGESWQDFLNKAKSFRALLRYRDGAKVHLGTRPDEKQYDVTETOLREM 120

QY 121 CQOGFSLNALYALSAGVHFTGLGCVLEDOEHQVAKKERETPTTDSMPPLLRQAIEFLDHQ 180

Db 121 TENGFSLRDLGLYALSASVSHFTLGAVLEQOEHTAALDRPAAPDENLPLLRQAIEFLDHQ 180

QY 181 GAEPAPFLGLELICGLEKOL 201

Db 181 DGEQAFHLGLESIRGFEVOL 201

RESULT 4

A26948

C:tetracycline resistance protein - Haemophilus parainfluenzae

C:Species: Haemophilus parainfluenzae

C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 12-Jun-1998

C:Accession: A26948

R:Heuer, C.; Hickman, R.K.; Curiale, M.S.; Hillen, W.; Levy, S.B.

J. Bacteriol. 169, 990-994, 1987

A:Title: Constitutive expression of tetracycline resistance mediated by a Tn10-like element

A:Reference number: A26948; MUID:87137315

A:Accession: A26948

A:Molecule type: DNA

A:Residues: 1-126 <HEU>

A:Cross-references: GB:M15539; NID:g148988; PID:g148989

C:Genetics:

A:Gene: tetr

C:Superfamily: tetracycline repressor

F:25-46/Region: helix-turn-helix #status predicted

Query Match 57.3%; Score 619; DB 2; Length 126;
Best Local Similarity 95.2%; Pred. No. 1.7e-50;
Matches 119; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML 60

Db 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML 60

QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

Db 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

QY 121 CQOGF 125

Db 121 ANKVF 125

RESULT 5

RPECYS

tetracycline repressor - Escherichia coli plasmid pSC101

C:Species: Escherichia coli

C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 12-Jun-1998

C:Accession: A03575

R:Brow, M.A.D.; Pesin, R.; Sutcliffe, J.G.

Mol. Biol. Evol. 2, 1-12, 1985

A:Title: The tetracycline repressor of pSC101.

A:Reference number: A03575; MUID:88216101

A:Accession: A03575

A:Molecule type: DNA

A:Residues: 1-219 <BRO>

A:Cross-references: GB:M36272; NID:g150945; PID:g150946

C:Comment: This protein is the repressor of the tetracycline resistance element; its

C:Genetics:

A:Gene: tetr

A:Genome: plasmid

C:Superfamily: tetracycline repressor

C:Keywords: antibiotic resistance; DNA binding; transcription regulation

F:24-46/Domain: DNA binding #status predicted <DBN>

F:25-46/Region: helix-turn-helix #status predicted

Query Match 48.6%; Score 525; DB 1; Length 219;
Best Local Similarity 52.7%; Pred. No. 1.7e-41;
Matches 107; Conservative 32; Mismatches 60; Indels 4; Gaps 1;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML 60

Db 1 MNKLOEAVIRTALELLNDVGMGLTRRLAERLGVQPPALYWHFKNKRALDLALAEML 60

61	QY	DRHHTFCPLGEGSWODFLNKA	KSPRCALLSHROGAKVHLGTR	TEKQYETLENOLAFL	120
61	DB	TIHHTTTPDDDDWSRFLKGN	CSFRALLAYRGARHAGTR	PAAPOMEXADQAQLAFL	120
121	QY	COQGFLENALYALSVA	HGHTLCGLVEDQEHQVAKERE	-----TPTTDSMPLLKQAEL	176
121	DB	CDAGFSAGDATYALMAIS	YTVGAVLEEQQA	SEADAERGEDOLTTTSASTMPARLOSAMKI	180
177	QY	FDHQGAEPAPFLFGL	ELIICGLEK	199	
181	DB	VIYEGGPDAAEP	GLIILIGLEK	203	

RESULT 6

 JQ1478
 regulatory protein tetr - Escherichia coli transposon Tn1721
 N:Alternate names: tetracycline resistance repressor
 C:Species: Escherichia coli
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 12-Jun-1998
 C:Accession: JQ1478; S02668; S24113
 R:Almeier, H.; Crenan, B.; Greck, M.; Schmitt, R.
 Gene 111, 11-20, 1992

 A:Title: Complete nucleotide sequence of Tn1721: gene organization and a novel gene product
 A:Reference number: JQ1475; MUID:92192465
 A:Accession: JQ1478
 A:Molecule type: DNA
 A:Residues: 1-216 <ALL>
 A:Cross-references: EMBL:X61367; NID:g48194; PID:g48198
 A:Note: the authors translated the codon GAG for residue 19 as Asp and CAG for residue 88
 R:Altschmid, L.; Baumeister, R.; Pfeleiderer, K.; Hillen, W.
 EMBO J. 7, 4011-4017, 1988

 A:Title: A threonine to alanine exchange at position 40 of tet repressor alters the repressor
 A:Reference number: S02667; MUID:89091153
 A:Accession: S02668
 A:Status: not compared with conceptual translation

 A:Molecule type: DNA
 A:Residues: 1-100 <ALT>
 C:Genetics:
 A:Gene: tetR
 C:Superfamily: tetracycline repressor
 F:25-46/Region: helix-turn-helix #status predicted

[illegible]

```

RESULT 7
S38655
tetr protein - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 12-Jun-1998
C:Accession: S38655
R:Trueman, P.; Sharpe, G.S.; Barth, P.T.
submitted to the EMBL Data Library, November 1993

```

A: Reference number: S30655
A: Accession: X30655
A: Status: Partially sequenced
A: Status: Partially sequenced
A: Residues: 216 <R>
A: Cross-references: EMBL:X75761; ND:g415984; PID:g415985
C: Superfamily: tetracycline repressor
E: 25-454/Region: helix-turn-helix <status predicted>

[illegible]

```

RESULT      8
I77569
tet repressor - Escherichia coli
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 12-Jun-1998
C:Accession: I77569
C:Toovar, K.; Ernst, A.; Hillen, W.
Mol. Gen. Genet. 215, 76-80, 1988
A:Title: Identification and nucleotide sequence of the class E tet regulatory element
A:Reference number: I57766; MUID:89201249
A:Accession: I77569
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: EMBL:X14035; NID:g43055; PID:g43056
C:Superfamily: tetracycline repressor
F:25-46/Region: helix-turn-helix #status predicted

```

```

Query Match      47.2%  Score 509.5;  DB 2;  Length 211;
Best Local Similarity 51.2%;  Pred. No. 4.4e-40;
Matches 103;  Conservative 29;  Mismatches 66;  Indels 3;  Gaps 1;

QY      1  MSRLDKSVINSALELLNEVGIEGLTRTKLAQKLGVEQPTLYVHWKKNRALLDALAIEML 60
DB      1  MARLSUDDVYSNAFLTLDSEGLEGLTRTKLAQSLKIEOPTLYVHWNRKQOTLMNMLSEAL 60

QY      61  DRHHHTFCPLEGESNQDFLRNKAASFRCALLSHRDGAKVHLGTRPTKEKYETLENQALFL 120
DB      61  AKHHTRSAPLPTESSWQOQLQNALSFRRKALVLRHDGARLHGHSPTPPDFEQEAQLRCL 120

QY      121  CQOGFSLENALYALSVAHGFTIGCVLEDEQEHQVAKERETPTTDSMPPLLRQALFIEDHQ 180
DB      121  CDAGFSVEEALFIQSISHTLGALEE--QATNQIENNHNVDAAAPPLQLEAFNIQART 177

QY      181  GAEPAPFLGELIILICGLEKQL 201
DB      178  SAEMAFHFGKSLIFGFSQAOL 198

RESULT      9
RPECRI

```

tet repressor protein (Tn 1721) - Escherichia coli plasmid RP1
 C:Species: Escherichia coli
 C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 12-Jun-1998
 C:Accession: A03574
 R:Waters, S.H.; Rogowsky, P.; Grinstead, J.; Altenbuchner, J.; Schmitt, R.
 Nucleic Acids Res. 11, 6089-6105, 1983
 A:Title: The tetracycline resistance determinants of RP1 and Tn1721: nucleotide sequence
 A:Reference number: A93486; MUID:83299270
 A:Accession: A03574
 A:Molecule type: DNA
 A:Residues: 1-216 <WAT>
 A:Cross-references: GB:X00006; NID:g42508; PID:g42509
 C:Genetics:
 A:Gene: tetR
 A:Genome: plasmid
 C:Superfamily: tetracycline repressor
 C:Keywords: antibiotic resistance; DNA binding; transcription regulation
 F:25-46/Region: helix-turn-helix #status predicted

Query Match 46.8%; Score 505.5; DB 1; Length 216;
 Best Local Similarity 51.5%; Pred. No. 1.1e-39;
 Matches 104; Conservative 34; Mismatches 61; Indels 3; Gaps 2;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLVGEQPTLYWHVKNKRALDIALAIEML 60
 Db 1 MTKLQPTVIRAAIDLLNEVGVDGLTRKLAERLVGQPPALYWHFNRKRALDIALAIEML 60
 QY 61 DRHHTHFCPLEGESWQDFLRNKAFCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120
 Db 61 AENHTSVPRADDDWRSFELTGNARFROALLAYRDGARIHAGTRPGAPOMETADAQLRFL 120
 QY 121 COGFSLENALYSAGVHFTLGCVLEQDQEHQVAKEREETPTDTSMP--PLLROAIELED 178
 Db 121 CEAGFSAGDAVNALMTISYFTVGAVLEQAGDSSEGER-GCTVEQAPLSPLLRADDAFD 179
 QY 179 HOGAEPALFGLLEILICLEKQ 200
 Db 180 EAGPDPAEQGLAVIVDGLAKR 201

RESULT 10
 H70740
 Hypothetical protein Rv1353c - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998
 C:Accession: H70740
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nature 393, 537-544, 1998
 A:Authors: Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, J.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: H70740
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-261 <COL>
 A:Cross-references: GB:275555; GB:AL123456; NID:g3261608; PID:e250355; PID:g1419056
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: Rv1353c

Query Match 16.3%; Score 176.5; DB 2; Length 261;
 Best Local Similarity 28.1%; Pred. No. 3.7e-09;
 Matches 61; Conservative 31; Mismatches 88; Indels 37; Gaps 5;

QY 4 LDKSVINSALELLNEVGIEGLTRKLAQKLVGEQPTLYWHVKNKRALDIALAIEMLDRH 63
 Db 16 INPEDIISGAFELAQQVSDINLSMPLGLHGVGVTSITWYFRKDDLLNAMTDRALSKY 75
 QY 64 HTFCPLEGESWQDFLRNKAFCR-----CALLSHRDGAKVHLGTRPTKEQYETLE 114

Db 76 VFATPYTEAGDWRETLNRHARSMRKTFADNPVLCDDLIR-----AALSPKTAIRLGAQEME 131
 QY 115 NQALFLCQGFSLLENALYSAGVHFTLGCVLEQDQ-----EHQVAKERE 159
 Db 132 KAIANLVTAGLSLEDAFDIYSAYSVHVRGSLVLDRLSRKSQSAGSGPSAIEHPVAIDPAT 191
 QY 160 TPTTDSMPPLLRQAIELFDHOGA--EPAPFLGLELII 194
 Db 192 T-----PLLAHATGGRHGRIGAPDETINFEYGLECIL 221

RESULT 11
 A40046
 Tet repressor homolog actII-1 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 12-Jun-1998
 C:Accession: A40046
 R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.
 Cell 66, 769-780, 1991
 A:Title: The act cluster contains regulatory and antibiotic export genes, direct target of the actA gene
 A:Reference number: A40046; MUID:91347376
 A:Accession: A40046
 A:Molecule type: DNA
 A:Residues: 1-259 <PER>
 A:Cross-references: GB:M64683; NID:g153143; PID:g455360
 C:Genetics:
 A:Start codon: GTG
 F:52-73/Region: helix-turn-helix #status predicted

Query Match 13.8%; Score 149.5; DB 2; Length 259;
 Best Local Similarity 26.0%; Pred. No. 1.1e-06;
 Matches 57; Conservative 35; Mismatches 90; Indels 37; Gaps 8;

QY 4 LDKSVINSALELLNEVGIEGLTRKLAQKLVGEQPTLYWHVKNKRALDIAL-----AI 57
 Db 31 LQDRIIVTALGIDAEGLDLSMRRLAQLAGTGHSAHYAHVGNRDELDFVDFIVLTV 90
 QY 58 EMLDRHHTHFCPLEGESWQDFLRNKAFCALLSHRDGAKVHLGTRPT-KEQYETLENQ 116
 Db 91 EYVE-----PEPGR-WAEQVKEMCRSLRMFLAHRDLARIAIDRVPLGPNMGVGMERT 142
 QY 117 LAFLCQGFSLLENALYSAGVHFTLGCVLEQDQEHQVAKERE-----RET 160
 Db 143 MNLRSGLHDELAAAYGGDLSTFTVALEQSSRNPTGEQAGVFADQLHGYLKS 202
 QY 161 PTTDSMPPLLRQA--IELFDHOGAEPAPFLGLELIIIGL 197
 Db 203 PAT-SFPNVLHLAGPIISLD---SDRRFELGLEIIAGL 237

RESULT 12
 A69900
 Hypothetical protein yobS - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1998
 C:Accession: A69900
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Gh
 wood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M
 Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Laube
 A:Authors: Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mael, C.; Med
 , K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Por
 Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.; Sch
 A:Authors: Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Ser
 amakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vanden
 ; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikaw
 A:Authors: Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PID:e1184426; PID:g2636046
A:Experimental source: strain 168
C:Genetics:
A:Gene: yvkB

```

Query Match      10.0%; Score 108.5; DB 2; Length 189;
Best Local Similarity 19.6%; Pred. No. 0.0047;
Matches 41; Conservative 32; Mismatches 51; Indels 85; Gaps 5;
QY 3 RLDKSKVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALIDALAIEMLDR 62
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 RPTNKRILDAAMQLLVKGYRATTKEIAEKANVSEATIFRNFKNKQGLVEALLSOHSSN 51
QY 63 HHTHFCPLLEGESWODEFLRNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFLCQ 122
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 RGSILEQTEGDLKDLL-----HIGT----- 83
QY 123 QGFSLENALYALS AVGHFTLGCVLEDOHQ-----VAKEERETPTTDSM 166
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
167 PPLLROAIELFDHOG-----AEPALFLF 188
QY 120 KQLLYDYLTGMEKGVITGNEAEHADVF 148
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: June 9, 1999, 12:59:50
Job time: 9169 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 10:29:01 ; Search time 16.52 Seconds
(without alignments)
336.316 Million cell updates/sec

Title: US-08-486-814-17
Perfect score: 1080
Sequence: 1 MSRLDKSKVINSALELNEV.....FGLELICGLEKOLKCESGS 207

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt_36:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1075	99.5	207	1	TER2_ECOLI	P04483 escherichia
2	683	63.2	218	1	TER8_PASPI	P51562 pasteurella
3	678	62.8	217	1	TER4_ECOLI	P09164 escherichia
4	673	62.3	207	1	TER8_PASMU	P51561 pasteurella
5	525	48.6	219	1	TER3_ECOLI	P03039 escherichia
6	515.5	47.7	216	1	TER1_ECOLI	P03038 escherichia
7	509.5	47.2	211	1	TER5_ECOLI	P21337 escherichia
8	487	45.1	210	1	TER7_VIBAN	P51560 vibrio angu
9	176.5	16.3	261	1	Y07H_MYCTO	Q11023 mycobacteri
10	110.5	10.2	210	1	MTRR_NEIGO	P39897 neisseria g
11	95.5	8.8	196	1	UIDR_ECOLI	Q59431 escherichia
12	85.5	7.9	677	1	SKD3_MOUSE	Q60649 mus musculu
13	85.5	7.9	236	1	YCFQ_ECOLI	P27598 escherichia
14	84	7.8	188	1	YP23_STAAN	P23217 staphylococ
15	83	7.7	195	1	BET1_ECOLI	P17446 escherichia
16	83	7.7	519	1	NIFL_AZOVI	P30663 azotobacter
17	83	7.7	955	1	PHSL_IPOBA	P27598 ipomoea bat
18	82	7.6	260	1	YFAX_ECOLI	P77732 escherichia
19	81.5	7.5	299	1	YWFK_BACSU	P39647 bacillus su
20	80.5	7.5	577	1	THR1_SCHPO	Q09684 schizosacch
21	80	7.4	2133	1	FAB_PIG	P12263 sus scrofa
22	80	7.4	212	1	TTK_ECOLI	P06969 escherichia
23	79	7.3	192	1	BM3R_BACME	P43506 bacillus me
24	79	7.3	866	1	MISP_SCHJA	Q05870 schistosoma
25	78.5	7.3	782	1	VACB_HAEIN	P44907 haemophilus
26	78	7.2	215	1	ACRE_ECOLI	P34000 escherichia
27	77.5	7.2	2472	1	SPCN_HUMAN	Q13813 homo sapien
28	77.5	7.2	1448	1	TRC5_ECOLI	P27190 escherichia
29	77.5	7.2	212	1	YCDC_ECOLI	P75899 escherichia
30	77	7.1	2319	1	FAB_MOUSE	Q06194 mus musculu
31	77	7.1	1957	1	YD85_SCHPO	Q10411 schizosacch
32	76.5	7.1	220	1	ENVR_ECOLI	P31676 escherichia
33	76.5	7.1	723	1	HS9A_CHICK	P11501 gallus gall
34	76.5	7.1	963	1	KINH_HUMAN	P33176 homo sapien
35	76.5	7.1	1208	1	YDH6_SCHPO	Q92351 schizosacch
36	76	7.0	218	1	TTK_HAEIN	P29280 haemophilus
37	75	7.0	187	1	Y893_HAEIN	P44923 haemophilus
38	75.5	7.0	549	1	CF1A_DROME	P16241 drosophila
39	75.5	7.0	723	1	HS9B_HUMAN	P08238 homo sapien
40	75.5	7.0	866	1	MISP_SCHMA	P06198 schistosoma
41	75.5	7.0	1391	1	N157_YEAST	P40064 saccharomyc
42	75.5	7.0	888	1	UFO_MOUSE	Q00993 mus musculu
43	74.5	6.9	557	1	ATKA_ECOLI	P03959 escherichia

ALIGNMENTS

RESULT 1	TER2_ECOLI	STANDARD:	PRT:	207 AA.
ID	TER2_ECOLI			
AC	P04483;			
DT	13-AUG-1987 (REL. 05, CREATED)			
DT	13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	TETRACYCLINE REPRESSOR PROTEIN CLASS B (TRANSPONSON TN10).			
GN	TETR.			
OS	ESCHERICHIA COLI.			
OC	PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;			
OC	ENTEROBACTERIACEAE.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 84247342.			
RA	POSTLE K., NGUYEN T.T., BERTRAND K.P.;			
RL	NUCLEIC ACIDS RES. 12:4849-4863(1984).			
RN	[2]			
RP	SEQUENCE OF 1-22 FROM N.A.			
RX	MEDLINE; 84005886.			
RA	BERTRAND K.P., POSTLE K., WRAY L.V. JR., REZNIKOFF W.S.;			
RL	GENE 23:149-156(1983).			
RN	[3]			
RP	MUTAGENESIS.			
RX	MEDLINE; 89094871.			
RA	SMITH L.D., BERTRAND K.P.;			
RL	J. MOL. BIOL. 203:949-953(1988).			
CC	-1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE			
CC	ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX			
CC	STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES			
CC	THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE			
CC	(TETA) PROMOTER OPERATOR SITES.			
CC	-1- SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL			
CC	REGULATORS.			
DR	EMBL; X00694; G43052; -.			
DR	EMBL; J01830; G154846; -.			
DR	PIR; A03576; RPECTN.			
DR	PROSITE; PS01081; HTH_TETR_FAMILY; 1.			
KW	TRANSCRIPTION REGULATION; REPRESSOR; TRANSPOSABLE ELEMENT;			
KW	DNA-BINDING; ANTIBIOTIC RESISTANCE.			
FT	DNA_BIND 26 45			
FT	SITE 64 64			
FT	FT SITE 64 64			
FT	FT METAL 100 100			
FT	FT MUTAGEN 64 64			
FT	FT MUTAGEN 82 82			
FT	FT MUTAGEN 103 103			
FT	FT SEQUENCE 207 AA; 23354 MW; 4AB4DD5A CRC32;			
SQ				

Query Match 99.5%; Score 1075; DB 1; Length 207;
Best Local Similarity 99.5%; Pred. No. 5.6e-93;
Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	1	MSRLDKSKVINSALELNEVEGLETTTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML	60
DB	1	MSRLDKSKVINSALELNEVEGLETTTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML	60
OY	61	DRHHTHFCPLEGESQWDFLRNNAKSKFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL	120
DB	61	DRHHTHFCPLEGESQWDFLRNNAKSKFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL	120

P12758 escherichia
P16749 human cytom

QY 121 COQGSFLENALYALSVAHGHTGCVLEQDEHQVAKEREETPTTDSMPPLLRQALIEFDHQ 180
 DB 121 COQGSFLENALYALSVAHGHTGCVLEQDEHQVAKEREETPTTDSMPPLLRQALIEFDHQ 180
 QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 2
 TER8_PASPI
 ID TER8_PASPI STANDARD; PRT; 218 AA.
 AC PS1562;
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DE 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 EN 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 GN TETRACYCLINE REPRESSOR PROTEIN CLASS H.
 OS PASTURELLA PISCICIDA.
 OG PLASMID P9351.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 CC PASTURELLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94328963.
 RA KIM E.H., AOKI T.;
 RL MICROBIOL. IMMUNOL. 38:31-38(1994).
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
 CC (TETA) PROMOTER OPERATOR SITES.
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC EMBL; D16172; G575938; -.
 DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;
 KW ANTIBIOTIC RESISTANCE; PLASMID.
 FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
 FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
 FT METAL 100 100 MAGNESIUM) OF [MG-TC]+ COMPLEX) (BY
 FT SIMILARITY).
 SQ SEQUENCE 218 AA; 24419 MW; 3C57D4AB CRC32;

Query Match 63.2%; Score 683; DB 1; Length 218;
 Best Local Similarity 65.2%; Pred. No. 1.5e-56;
 Matches 131; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 60
 DB 1 MARLNRESVIDAALLNETGIDGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 60
 QY 61 DRHHTHFCPLEGSWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQAF 120
 DB 61 ARHHDYSLPAAGESWQSFRLNNAQKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQAF 120
 QY 121 COQGSFLENALYALSVAHGHTGCVLEQDEHQVAKEREETPTTDSMPPLLRQALIEFDHQ 180
 DB 121 TENGFSRLDGLYALSVAHSHFTLGAVLEQDEHQVAKEREETPTTDSMPPLLRQALIEFDHQ 180
 QY 181 GAEPAPFLGLELIICGLEKQK 201
 DB 181 DGEQAFHGLGSLIRGFEVOL 201

RESULT 3
 TER4_ECOLI
 ID TER4_ECOLI STANDARD; PRT; 217 AA.
 AC P09164;
 DT 01-MAR-1989 (REL. 10, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS D.
 GN TETR.
 OS ESCHERICHIA COLI, AND SALMONELLA ORDONEZ.
 OG PLASMID PIP173.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 CC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-E.COLI; STRAIN-J53-1; PLASMID-RA1;
 RX MEDLINE; 85037938.
 RA UNGER B., KLOCK G., HILLEN W.;
 RL NUCLEIC ACIDS RES. 12:7693-7703(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S.ORDONEZ; STRAIN-BM2000; PLASMID-PIP173;
 RX MEDLINE; 93204906.
 RA ALLARD J.D., GIBSON M.L., VU L.H., NGUYEN T.T., BERTRAND K.P.;
 RL MOL. GEN. GENET. 237:301-305(1993).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS), AND PARTIAL SEQUENCE.
 RX MEDLINE; 94204640.
 RA HINRICHS W., KISKER C., DUEVEL C., MUELLER A., TOVAR K., HILLEN W.,
 RA SAENGER W.;
 RL SCIENCE 264:418-420(1994).
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
 CC (TETA) PROMOTER OPERATOR SITES.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC EMBL; X01083; G43054; -.
 DR EMBL; X65876; G49075; -.
 DR PIR; S07359; S07359.
 DR PIR; S30287; S30287.
 DR PDB; 2TCT; 03-APR-96.
 DR PDB; 2TRT; 20-JUN-96.
 DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;
 KW ANTIBIOTIC RESISTANCE; 3D-STRUCTURE.
 FT INIT_MET 0 0
 FT DNA_BIND 25 44 H-T-H MOTIF (POTENTIAL).
 FT SITE 63 63 INVOLVED IN BINDING TO [MG-TC]+.
 FT METAL 99 99 MAGNESIUM (OF [MG-TC]+ COMPLEX).
 SQ SEQUENCE 217 AA; 24287 MW; 830EF8EB CRC32;

Query Match 62.8%; Score 678; DB 1; Length 217;
 Best Local Similarity 65.0%; Pred. No. 4.4e-56;
 Matches 130; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 2 SRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 61
 DB 1 ARLNRESVIDAALLNETGIDGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 60
 QY 62 RHHTHFCPLEGSWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQAF 121
 DB 61 RHHDYSLPAAGESWQSFRLNNAQKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQAF 120
 QY 122 COQGSFLENALYALSVAHGHTGCVLEQDEHQVAKEREETPTTDSMPPLLRQALIEFDHQ 181
 DB 122 ENGFSRLDGLYALSVAHSHFTLGAVLEQDEHQVAKEREETPTTDSMPPLLRQALIEFDHQ 181
 QY 182 AEPAPFLGLELIICGLEKQK 201
 DB 181 GEPAPFLGLELIICGLEKQK 200

RESULT 4


```

TER8_PASMU          STANDARD;          PRT;    207 AA.
ID  TER8_PASMU
AC  P51561;
DT  01-OCT-1996 (REL. 34, CREATED)
DT  01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  TETRACYCLINE REPRESSOR PROTEIN CLASS H.
GN  TETR.
OS  PASTURELLA MULTOCIDA.
OG  PLASMID PVML11.
OC  PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC  PASTURELLACEAE.
RN  [1]
RC  SEQUENCE FROM N.A.
RC  STRAIN=P2862;
RX  MEDLINE; 94153051.
RA  HANSEN L.M., MCMURRY L.M., LEVY S.B., HIRSH D.C.;
RL  ANTIMICROB. AGENTS CHEMOTHER. 37:2699-2705(1993).
CC  -1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC  ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC  STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC  THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC  (TETR) PROMOTER OPERATOR SITES.
CC  -1- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
CC  -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC  REGULATORS.
DR  EMBL; U00792; G392872; -.
DR  PROSITE; PS01081; HTH_TETR_FAMILY; 1.
KW  TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;
KW  ANTI-BIOTIC RESISTANCE; PLASMID.
FT  DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT  SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT  METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
FT  SEQUENCE 207 AA; 23156 MW; A51DA28B CRC32;

Query Match 62.3%; Score 673; DB 1; Length 207;
Best Local Similarity 64.2%; Pred. No. 1.2e-55;
Matches 131; Conservative 27; Mismatches 44; Indels 2; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
D 1 MAKLDKEQVIDDALILLNEVGIEGLTRRNVAQKIGVEQPTLYWHVKNRALLDALAETIL 60
QY 61 DRHHTHFCPLEGESWQDFLRNKAQKLSHRDGAQKVLHGTPTKQYETLENQALFL 120
D 61 QKHHVLPPLNETWQDFLRNKAQKLSHRDGAQKVLHGTPTKQYETLENQALFL 120
QY 121 COQGSFLENALYASAVGHFTLGCVLDEQHOVAKEEREPTTDSM--PPLLRQAIELFD 178
D 121 CDAGFSLSQAVVALSSIAHFTLGSVLTEQHQESQKEREKVTDTVAYPPLLTQAVAIMD 180
QY 179 HOGAEPALFGLLELICGLEKQLK 202
D 181 SDNGDAALFLVLDVMISGLETVLK 204

RESULT 5
TER3_ECOLI
ID TER3_ECOLI          STANDARD;          PRT;    219 AA.
AC  P03039;
DT  21-JUL-1986 (REL. 01, CREATED)
DT  21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DE  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  TETRACYCLINE REPRESSOR PROTEIN CLASS C.
GN  TETR.
OS  ESCHERICHIA COLI.
OG  PLASMID PSC101.
OC  PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC  ENTEROBACTERIACEAE.
RN  [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE; 88216101.
RA BROW M.A.D., PESIN R., SUTCLIFFE J.G.;
RL MOL. BIOL. EVOL. 2:11-12(1985).
CC -1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETR) PROMOTER OPERATOR SITES.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; M36272; G150946; -.
DR PIR; A03575; RPECYS.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;
KW ANTI-BIOTIC RESISTANCE.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
FT SEQUENCE 219 AA; 24174 MW; 6C9E1828 CRC32;

Query Match 48.6%; Score 525; DB 1; Length 219;
Best Local Similarity 52.7%; Pred. No. 7.2e-42;
Matches 107; Conservative 32; Mismatches 60; Indels 4; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
D 1 MNKLQREAVIRTALELLNDVGMGLTTRRLAERLGVQOPALYWHFKNRALLDALAEAML 60
QY 61 DRHHTHFCPLEGESWQDFLRNKAQKLSHRDGAQKVLHGTPTKQYETLENQALFL 120
D 61 TINHTHSTPRDDDDWRSFLKGNACSFRRALLAYRDGARIHAGTRPAAPQMEKADQLRFL 120
QY 121 COQGSFLENALYASAVGHFTLGCVLDEQHOVAKEERE---TPTDSMPPLLRQAIEL 176
D 121 CDAGFSAGDATYALMAISYFTVGAVLEQOASEADAEERGEDQLTTSASTMPARLQSAWKI 180
QY 177 FDHQAEPALFGLLELICGLEK 199
D 181 VVEGGDAAAFERGLALIIGGLEK 203

RESULT 6
TER1_ECOLI
ID TER1_ECOLI          STANDARD;          PRT;    216 AA.
AC  P03038;
DT  21-JUL-1986 (REL. 01, CREATED)
DT  01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  TETRACYCLINE REPRESSOR PROTEIN CLASS A (TRANSPONSON 1721).
GN  TETR.
OS  ESCHERICHIA COLI.
OG  PLASMID RP1, AND PLASMID RP4.
OC  PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC  ENTEROBACTERIACEAE.
RN  [1]
RC  SEQUENCE FROM N.A.
RC  TRANSPONSON-TN1721;
RX MEDLINE; 92192465.
RA ALLMEIER H., CRESNAR B., GRECK M., SCHMITT R.;
RL GENE 111:11-20(1992).
RN [2]
RC  SEQUENCE FROM N.A.
RC  PLASMID-PLASMID RP4;
RA TRUEMAN P., SHARPE G.S., BARTH P.T.;
RL SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RC  SEQUENCE FROM N.A.
RC  TRANSPONSON-TN1721; PLASMID-RP1;
RX MEDLINE; 83299270.

```

RA WATERS S.H., ROGOWSKY P., GRINSTED J., ALTENBUCHNER J., SCHMITT R.;
 RL NUCLEIC ACIDS RES. 11:6089-6105(1983).
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
 CC (TETA) PROMOTER OPERATOR SITES.
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC EMBL; X61367; G48198; -.
 DR EMBL; X75761; G415985; -.
 DR EMBL; X00006; G42509; -.
 DR PIR; A03574; RPECRI.
 DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
 KW TRANSCRIPTION REGULATION; REPRESSOR; TRANSPOSABLE ELEMENT;
 KW DNA-BINDING; ANTIBIOTIC RESISTANCE; PLASMID.
 FT DNA_BIND 26 45
 FT SITE 64 64
 FT INVOLVED IN BINDING TO [MG-TC]+ (BY
 FT SIMILARITY).
 FT METAL 100 100
 FT MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
 FT SIMILARITY).
 FT TH -> ST (IN REF. 3).
 FT I -> T (IN REF. 3).
 FT DA -> ES (IN REF. 3).
 FT SEQUENCE 216 AA; 23320 MW; F72BDFB1 CRC32;

Query Match 47.7%; Score 515.5; DB 1; Length 216;
 Best Local Similarity 52.5%; Pred. No. 5.4e-41;
 Matches 106; Conservative 33; Mismatches 60; Indels 3; Gaps 2;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAOKLGVQPTLYWHVKNKRALDLALAIEM 60
 DB 1 MTKLPNTVIRALDNEVGVDGLTRKLAERLGVQVQPALYHFRNKRALLDALAEAML 60
 QY 61 DRHHTHFCPLGESQDPLRNKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQLAFL 120
 DB 61 AENHTSVPRADDDWSRFLGNARSFRQALLAYRDGARIHAGTRPGAPQMETADAQLRFL 120
 QY 121 COOGFSLENALYALSAGVHFTLGCYLEDQEHQVAKEREETPTTDSMP--PILRQAIELFD 178
 DB 121 CEAGFSAGDAVNALMTSYFTVGVAVLEQAGSDSAGER-GGTVEQAPLPLRAAIDAFD 179
 QY 179 HOGAEPAPFLFGLLEIICGLEKQ 200
 DB 180 EAGPDAFAEQGLAVIVDGLAKR 201

RESULT 7
 TERS_ECOLI
 ID TERS_ECOLI STANDARD; PRT; 211 AA.
 AC P21337;
 DT 01-MAY-1991 (REL. 18, CREATED)
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)
 WJ 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS E.
 GN TETR.
 OS ESCHERICHIA COLI.
 OG PLASMID PS11503.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21.
 RX MEDLINE; 89201249.
 RA TOVAR K., ERNST A., HILLEN W.;
 RA MOL. GEN. GENET. 215:76-80(1988).
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
 CC (TETA) PROMOTER OPERATOR SITES.
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL

CC REGULATORS.
 DR EMBL; M34933; G155020; ALT_SEQ.
 DR EMBL; X14035; G43056; -.
 DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;
 KW ANTIBIOTIC RESISTANCE; PLASMID.
 FT DNA_BIND 26 45
 FT SITE 64 64
 FT INVOLVED IN BINDING TO [MG-TC]+ (BY
 FT SIMILARITY).
 FT METAL 100 100
 FT MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
 FT SIMILARITY).
 FT CONFLICT 45 45
 FT V -> L (IN G43056).
 SQ SEQUENCE 211 AA; 23585 MW; F538C813 CRC32;

Query Match 47.2%; Score 509.5; DB 1; Length 211;
 Best Local Similarity 51.2%; Pred. No. 1.9e-40;
 Matches 103; Conservative 29; Mismatches 66; Indels 3; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAOKLGVQPTLYWHVKNKRALDLALAIEM 60
 DB 1 MARLSLDDVISNALTLIDSEGLEGLTRKLAOSLAKIEQPTLYWHVKNKOTLNNMLSEAIL 60
 QY 61 DRHHTHFCPLGESQDPLRNKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQLAFL 120
 DB 61 AKHTRSAPLPTESWQOFLQENALSFRRKALLVHRDGRALHIGTSPPTPQFEQAQLRCL 120
 QY 121 COOGFSLENALYALSAGVHFTLGCYLEDQEHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180
 DB 121 CDAGFSVEALFISGHSHTLGAVLEE---OATNQIENNHVDAAPPLQAEAFNIQART 177
 QY 181 GAEPAPFLFGLLEIICGLEKQ 201
 DB 178 SAEMAFHFGKSLIFGFSQAL 198

RESULT 8
 TERS_VIBAN
 ID TERS_VIBAN STANDARD; PRT; 210 AA.
 AC P51560;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS G.
 GN TETR.
 OS VIBRIO ANGUILLARUM.
 OG PLASMID PJA8122.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC VIBRIONACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-M28122;
 RX MEDLINE; 93125225.
 RA ZHAO J., AOKI T.;
 RL MICROBIOL. IMMUNOL. 36:1051-1060(1992).
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
 CC (TETA) PROMOTER OPERATOR SITES.
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC EMBL; S52438; G262929; -.
 DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;
 KW ANTIBIOTIC RESISTANCE; PLASMID.
 FT DNA_BIND 26 45
 FT SITE 64 64
 FT INVOLVED IN BINDING TO [MG-TC]+ (BY
 FT SIMILARITY).
 FT METAL 100 100
 FT MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
 FT SIMILARITY).
 SQ SEQUENCE 210 AA; 23595 MW; 811CB332 CRC32;

Fri Jun 11 11:58:14 1999

```

DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DE 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
GN MTRR.
OS NEISSERIA GONORRHOEA.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC NEISSERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FA19;
RX MEDLINE; 94254732.
RA PAN W., SPRATT B.G.;
RL MOL. MICROBIOL. 11:769-775(1994).
CC -!- FUNCTION: PUTATIVE REPRESSOR OF MTRC GENE. CONTROLS THE
CC PERMEABILITY OF THE CELL ENVELOPE TO HYDROPHOBIC COMPOUNDS SUCH AS
CC ANTIBIOTICS AND DETERGENTS.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC EMBL: Z25797; G438189; -.
DR PIR; S40250; S40250.
DR PROSITE; PS01081; HTH_TETR_FAMILY: 1.
KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.
FT DNA_BIND 32 51 H-T-H MOTIF (POTENTIAL).
FT VARIANT 105 105 H -> Y; IN PENICILLIN-RESISTANT ISOLATES.
SQ SEQUENCE 210 AA; 24192 MW; F4CABF98 CRC32;

Query Match 10.2%; Score 110.5; DB 1; Length 210;
Best Local Similarity 27.4%; Pred. No. 0.0021;
Matches 52; Conservative 22; Mismatches 73; Indels 43; Gaps 8;

QY 6 KSKVINSALELLENEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEMLDHRHT 65
DB 12 KEHLMLALETFFYKKGARTSLNEIAQAAGVTRGALYWHKKNEDLFDALFORICDDIEN 71
QY 66 ----HFCPLEGESQWDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTKQYETLENQALFLC 121
DB 72 CIAQDAADAEGSWTVF-RHTLLHFFERLQSNIDHYKHF-----NILFLKC 116
QY 122 QQGSLENALYALSVAHGFTLGCVLDEQEHQVAKKEERTPTTDSMPPLLRQATE---LFD 178
DB 117 EH--TEQNA--AVTATA-----RKHOAIWREKIT-----AVLVEAVENQDLAD 155
QY 179 HOGAEPALFL 188
DB 156 DLDKETAVIF 165

RESULT 11
UIDR_ECOLI
ID UIDR_ECOLI STANDARD; PRT; 196 AA.
AC Q59431.
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE UID OPERON REPRESSOR (GUS OPERON REPRESSOR).
GN UIDR OR GUSR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA WILSON K.J.; JEFFERSON R.A.;
RL SUBMITTED (OCT-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA BLATTNER F.R.; PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.

```

```

DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DE 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
GN MTRR.
OS NEISSERIA GONORRHOEA.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC NEISSERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FA19;
RX MEDLINE; 94254732.
RA PAN W., SPRATT B.G.;
RL MOL. MICROBIOL. 11:769-775(1994).
CC -!- FUNCTION: PUTATIVE REPRESSOR OF MTRC GENE. CONTROLS THE
CC PERMEABILITY OF THE CELL ENVELOPE TO HYDROPHOBIC COMPOUNDS SUCH AS
CC ANTIBIOTICS AND DETERGENTS.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC EMBL: Z25797; G438189; -.
DR PIR; S40250; S40250.
DR PROSITE; PS01081; HTH_TETR_FAMILY: 1.
KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.
FT DNA_BIND 32 51 H-T-H MOTIF (POTENTIAL).
FT VARIANT 105 105 H -> Y; IN PENICILLIN-RESISTANT ISOLATES.
SQ SEQUENCE 210 AA; 24192 MW; F4CABF98 CRC32;

Query Match 10.2%; Score 110.5; DB 1; Length 210;
Best Local Similarity 27.4%; Pred. No. 0.0021;
Matches 52; Conservative 22; Mismatches 73; Indels 43; Gaps 8;

QY 6 KSKVINSALELLENEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEMLDHRHT 65
DB 12 KEHLMLALETFFYKKGARTSLNEIAQAAGVTRGALYWHKKNEDLFDALFORICDDIEN 71
QY 66 ----HFCPLEGESQWDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTKQYETLENQALFLC 121
DB 72 CIAQDAADAEGSWTVF-RHTLLHFFERLQSNIDHYKHF-----NILFLKC 116
QY 122 QQGSLENALYALSVAHGFTLGCVLDEQEHQVAKKEERTPTTDSMPPLLRQATE---LFD 178
DB 117 EH--TEQNA--AVTATA-----RKHOAIWREKIT-----AVLVEAVENQDLAD 155
QY 179 HOGAEPALFL 188
DB 156 DLDKETAVIF 165

RESULT 11
UIDR_ECOLI
ID UIDR_ECOLI STANDARD; PRT; 196 AA.
AC Q59431.
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE UID OPERON REPRESSOR (GUS OPERON REPRESSOR).
GN UIDR OR GUSR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA WILSON K.J.; JEFFERSON R.A.;
RL SUBMITTED (OCT-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA BLATTNER F.R.; PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.

```

STRAIN-K12;
 ALBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,
 IREMO K., INADA T., ISONO K., ISONO S., ITOH T., KANAI K., KASAI H.,
 KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M., KITAKAWA M., MAKINO K.,
 MASUDA S., MIKI T., MIZOBUCHI K., MORI H., MOTOMURA K., NAKAMURA Y.,
 NASHIMOTO H., NISHIO Y., OSHIMA T., SAITO N., SAMPEI G., SEKI Y.,
 TAGAMI H., TAKEMOTO K., WADA C., YAMAMOTO Y., YANO M.;
 SUBMITTED (DEC-1996) TO EMBL/GENBANK/DDJB DATA BANKS;
 -!- FUNCTION: REPRESSOR FOR THE UIDRABC (GUSRABC) OPERON.
 CC !- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL: M14641; G868019; -;
 DR EMBL: AE000257; G1787904; -;
 DR EMBL: J90805; G1742672; -;
 DR ECOGENE: EG12667; UIDR.
 DR PROSITE: PS01081; HTH_TETR_FAMILY; 1.
 KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.
 FT DNA_BIND 33 52
 SQ SEQUENCE 196 AA; 21799 MW; 6C1840FC CRC32;

Query Match	8.8%	Score 95.5;	DB 1;	Length 196;
Best Local Similarity	25.38;	Pred. No. 0.049,		
Matches	19;	Conservative 23;	Mismatches 32;	Indels 1; Gaps 1;
QY	6	KSQVNSALELNEVNGIEGTTTRKLAQKLGVEQFTLHVHKNRAILLDALATMLDRHHT	65	
Db	13	RTRIILNAAREIFSENGFHSASMKCAISCAISPTGLYHHFISKEALIQAILQOERALA	72	
QY	66	HFC-PLGESWQDFL	79	
Db	73	RRFPIEGIHFDVYM	87	

Db	523	FRRDEFLGRINEI-VYFLPCHSELQYNKLNFWAKRQKRHNITLWMDREAVDVLVD	581
QY	96	GAKVHGLTRPTEKOYE-TLENQAFICQQGFSENLAYLSAVGHFTTIGCVLEDOE-HQV	153
Db	582	GYNVHYGARSIKHEVERRVYNQAAAEQDL-----LPGCTLRITVEDSKHLL	631
QY	154	AKEREETPTDTSMPPLLRQAIELFD	178
Db	632	KSPQLSPQAEKRPPTIR--LEIID	654

RESULT	13
YCFO_ECOLI	
ID	YCFO_ECOLI STANDARD; PRT; 236 AA.
AC	P75952;
DC	15-JUL-1998 (REL. 36, CREATED)
DT	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DT	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN NDH-MFD INTERGENIC REGION.
GN	YCFO.
OS	ESCHERICHIA COLI.
OC	PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC	ENTEROBACTERIAEAE.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-K12 / MG1655;
RA	BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D. ;
RRL	SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CCC	- ! - SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.
DRR	EMBL; AEO00211; G1787354; ..
DDR	EUGENE; EG13435; YCFO.
DDR	PROSITE; PS01081; HTH_TETR_FAMILY; FALSE_NEG.
KKW	HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING.
DNA_BIND	67
SEQUENCE	236 AA; 26204 MW; AD1F64FC CRC32; H-T-H MOTIF (POTENTIAL).
TSQ	96

RESULT	14
P23 STAAU	
D D	VP23 STAAU STANDARD; PRT: 188 AA.
C C	P23217;
T T	01-NOV-1991 (REL. 20, CREATED)
T T	01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
E E	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
H H	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN QACA 5' REGION (ORF 188).
S S	STAPHYLOCOCCUS AUREUS.
G G	PLASMID PSK1.
C C	PROKARYOTA; COCCI; MICROCOCCACEAE.
N N	[1]
X X	SEQUENCE FROM N.A.
P P	MEDLINE; 91211614.
A A	ROUGH D.A., CRAM D.S., DIBERARDINO D., LITTLEJOHN T.G. SKURRAY R.A.;

```

Query Match          7.9%; Score 85.5; DB 1; Length 677;
Best Local Similarity 22.4%; Pred. No. 1.8;
Matches 46; Conservative 32; Mismatches 84; Indels 43; Gaps 10;

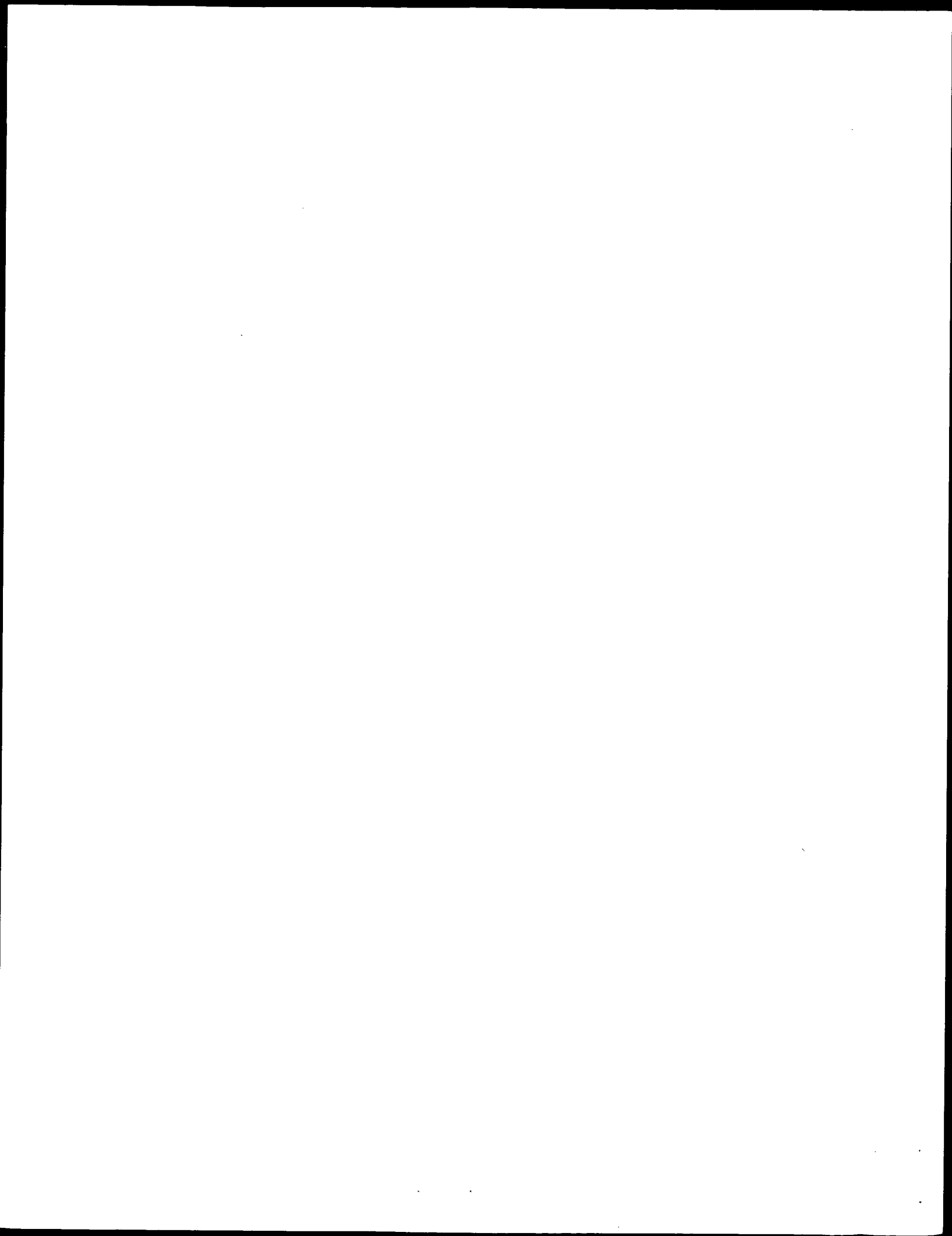
      Y      2 SRDKSKVINSALELNEVGIEGLTRTKAQLQK-----VEOPTLYWH 44
      | : : ||| : | : : : : ||
      b      465 SNVASDEIAQHAQLQROE-ALE-MSRNRTAENLGDVQMSDKRITTSKNFKENVIRPILKA 522
      Y      45 VKNKRAL--LDALATMLDRHHTFCPLLEGESQDFLNKAKSFRCALLSHR-----D 95

```

```

RESULT 15
ID BETI_ECOLI STANDARD; PRT; 195 AA.
AC P17446;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN BETI.
ES BETI.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RC MEDLINE; 92065800.
RX LAMARK T., KAASEN E., ESHOO M.W., FALKENBERG P., MCDUGALL J.,
RA STROM A.R.;
RM MOL. MICROBIOL. 5:1049-1064(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RC BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA DUNCAN M., ALLEN E., ARAUJO R., APARTICIO A.M., CHUNG E., DAVIS K.,
RA FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O., LEW H.,
RA LIN D., NAMATH A., OEFNER P., ROBERTS D., SCHRAMM S., DAVIS R.W.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -1- FUNCTION: REPRESSOR INVOLVED IN CHOLINE REGULATION OF THE BET
CC GENES.
CC -1- PATHWAY: OSMOREGULATORY CHOLINE-GLYCINE BETAINE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC EMBL; X52905; G48717; -.
CC EMBL; AE000138; G1786505; -.
CC EMBL; U73857; G1657511; ALT_INIT.
DR PIR; S10899; S10899.
DR PIR; S15180; S15180.
DR ECOGENE; EG10111; BETI.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
DR TRANSRIPTION REGULATION; DNA-BINDING; REPRESSOR; TRANS-ACTING FACTOR.

```



GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

```
run on:      June  9, 1999, 10:28:01 ; Search time 28.57 Seconds
              (without alignments)
              399.721 Million cell updates/sec
```

Title: US-08-486-814-17
 Perfect score: 1080
 Sequence: 1 MSRLDKSVINSALELLNEY.....FGLELICGLEQLKCESG 207

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database : SPTREMBL_8:*

```

1: sp_fungi.*
2: sp_human.*
3: sp_invertebrate.*
4: sp_mammal.*
5: sp_mmc.*
6: sp_organelle.*
7: sp_organism.*
8: sp_plant.*
9: sp_bacteria.*
10: sp_rodent.*
11: sp_virus.*
12: sp_vertebrate.*
13: sp_unclassified.*
14: sp_archaea.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	619	57.3	126	9	Q59466	Q59466	haemophilus
2	149.5	13.8	259	9	Q53901	Q53901	streptomyce
3	129.5	12.0	242	9	Q52558	Q52558	amycolatops
4	119	11.0	191	9	Q34892	Q34892	bacillus su
5	110.5	10.2	210	9	Q59608	Q59608	neisseria g
6	108.5	10.0	189	9	Q34643	Q34643	bacillus su
7	101.5	9.4	192	9	Q67927	Q67927	aquifex aeo
8	98	9.1	196	9	Q53789	Q53789	mycobacteri
9	94.5	8.8	186	9	Q51597	Q51597	pseudomonas
10	94.5	8.8	236	9	Q73186	Q73186	synechocyst
11	94	8.7	179	9	Q67157	Q67157	aquifex aeo
12	94	8.7	217	9	Q87854	Q87854	streptomyce
13	93.5	8.7	246	9	Q53295	Q53295	mycobacteri
14	92	8.5	213	9	Q53310	Q53310	mycobacteri
15	89	8.2	194	9	Q07001	Q07001	bacillus su
16	88.5	8.2	185	9	Q70020	Q70020	staphylococ
17	87.5	8.1	235	9	Q25951	Q25951	helicobacte
18	86.5	8.0	192	14	Q29994	Q29994	archaeoglob
19	86	8.0	205	9	Q51516	Q51516	pseudomonas
20	85	7.9	202	9	Q51730	Q51730	pseudomonas
21	85	7.9	198	9	P96676	P96676	bacillus su
22	84.5	7.8	2472	10	P70477	P70477	rattus norv
23	84.5	7.8	2472	10	Q88663	Q88663	rattus norv
24	84	7.8	207	9	Q34970	Q34970	bacillus su
25	84	7.8	188	9	Q59802	Q59802	staphylococ
26	83.5	7.7	435	2	Q75901	Q75901	homo sapien
27	83.5	7.7	216	9	P95100	P95100	mycobacteri
28	83.5	7.7	542	12	Q91290	Q91290	pleurodeles
29	83	7.7	405	3	Q18399	Q18399	rosophila

Q42863 ipomoea bat
Q07930 aquifex aeo
Q06919 azotobacter
Q13186 homo sapien
Q59147 pyrococcus
Q75181 homo sapien
Q54830 streptococc
Q54835 streptococc
Q08961 rattus norv
Q19127 macaca fasc
Q13126 macaca mula
Q60588 homo sapien
Q14812 homo sapien
Q53165 mycobacteri
Q59306 clostridium
Q07388 mycobacteri

ALIGNMENTS

RESULT 1

Q59466 PRELIMINARY; PRT; 126 AA.
 Q59467
 Q59468
 Q59469
 Q59470
 Q59471
 Q59472
 Q59473
 Q59474
 Q59475
 Q59476
 Q59477
 Q59478
 Q59479
 Q59480
 Q59481
 Q59482
 Q59483
 Q59484
 Q59485
 Q59486
 Q59487
 Q59488
 Q59489
 Q59490
 Q59491
 Q59492
 Q59493
 Q59494
 Q59495
 Q59496
 Q59497
 Q59498
 Q59499
 Q59500
 Q59501
 Q59502
 Q59503
 Q59504
 Q59505
 Q59506
 Q59507
 Q59508
 Q59509
 Q59510
 Q59511
 Q59512
 Q59513
 Q59514
 Q59515
 Q59516
 Q59517
 Q59518
 Q59519
 Q59520
 Q59521
 Q59522
 Q59523
 Q59524
 Q59525
 Q59526
 Q59527
 Q59528
 Q59529
 Q59530
 Q59531
 Q59532
 Q59533
 Q59534
 Q59535
 Q59536
 Q59537
 Q59538
 Q59539
 Q59540
 Q59541
 Q59542
 Q59543
 Q59544
 Q59545
 Q59546
 Q59547
 Q59548
 Q59549
 Q59550
 Q59551
 Q59552
 Q59553
 Q59554
 Q59555
 Q59556
 Q59557
 Q59558
 Q59559
 Q59560
 Q59561
 Q59562
 Q59563
 Q59564
 Q59565
 Q59566
 Q59567
 Q59568
 Q59569
 Q59570
 Q59571
 Q59572
 Q59573
 Q59574
 Q59575
 Q59576
 Q59577
 Q59578
 Q59579
 Q59580
 Q59581
 Q59582
 Q59583
 Q59584
 Q59585
 Q59586
 Q59587
 Q59588
 Q59589
 Q59590
 Q59591
 Q59592
 Q59593
 Q59594
 Q59595
 Q59596
 Q59597
 Q59598
 Q59599
 Q59600
 Q59601
 Q59602
 Q59603
 Q59604
 Q59605
 Q59606
 Q59607
 Q59608
 Q59609
 Q59610
 Q59611
 Q59612
 Q59613
 Q59614
 Q59615
 Q59616
 Q59617
 Q59618
 Q59619
 Q59620
 Q59621
 Q59622
 Q59623
 Q59624
 Q59625
 Q59626
 Q59627
 Q59628
 Q59629
 Q59630
 Q59631
 Q59632
 Q59633
 Q59634
 Q59635
 Q59636
 Q59637
 Q59638
 Q59639
 Q59640
 Q59641
 Q59642
 Q59643
 Q59644
 Q59645
 Q59646
 Q59647
 Q59648
 Q59649
 Q59650
 Q59651
 Q59652
 Q59653
 Q59654
 Q59655
 Q59656
 Q59657
 Q59658
 Q59659
 Q59660
 Q59661
 Q59662
 Q59663
 Q59664
 Q59665
 Q59666
 Q59667
 Q59668
 Q59669
 Q59670
 Q59671
 Q59672
 Q59673
 Q59674
 Q59675
 Q59676
 Q59677
 Q59678
 Q59679
 Q59680
 Q59681
 Q59682
 Q59683
 Q59684
 Q59685
 Q59686
 Q59687
 Q59688
 Q59689
 Q59690
 Q59691
 Q59692
 Q59693
 Q59694
 Q59695
 Q59696
 Q59697
 Q59698
 Q59699
 Q59700
 Q59701
 Q59702
 Q59703
 Q59704
 Q59705
 Q59706
 Q59707
 Q59708
 Q59709
 Q59710
 Q59711
 Q59712
 Q59713
 Q59714
 Q59715
 Q59716
 Q59717
 Q59718
 Q59719
 Q59720
 Q59721
 Q59722
 Q59723
 Q59724
 Q59725
 Q59726
 Q59727
 Q59728
 Q59729
 Q59730
 Q59731
 Q59732
 Q59733
 Q59734
 Q59735
 Q59736
 Q59737
 Q59738
 Q59739
 Q59740
 Q59741
 Q59742
 Q59743
 Q59744
 Q59745
 Q59746
 Q59747
 Q59748
 Q59749
 Q59750
 Q59751
 Q59752
 Q59753
 Q59754
 Q59755
 Q59756
 Q59757
 Q59758
 Q59759
 Q59760
 Q59761
 Q59762
 Q59763
 Q59764
 Q59765
 Q59766
 Q59767
 Q59768
 Q59769
 Q59770
 Q59771
 Q59772
 Q59773
 Q59774
 Q59775
 Q59776
 Q59777
 Q59778
 Q59779
 Q59780
 Q59781
 Q59782
 Q59783
 Q59784
 Q59785
 Q59786
 Q59787
 Q59788
 Q59789
 Q59790
 Q59791
 Q59792
 Q59793
 Q59794
 Q59795
 Q59796
 Q59797
 Q59798
 Q59799
 Q59800
 Q59801
 Q59802
 Q59803
 Q59804
 Q59805
 Q59806
 Q59807
 Q59808
 Q59809
 Q59810
 Q59811
 Q59812
 Q59813
 Q59814
 Q59815
 Q59816
 Q59817
 Q59818
 Q59819
 Q59820
 Q59821
 Q59822
 Q59823
 Q59824
 Q59825
 Q59826
 Q59827
 Q59828
 Q59829
 Q59830
 Q59831
 Q59832
 Q59833
 Q59834
 Q59835
 Q59836
 Q59837
 Q59838
 Q59839
 Q59840
 Q59841
 Q59842
 Q59843
 Q59844
 Q59845
 Q59846
 Q59847
 Q59848
 Q59849
 Q59850
 Q59851
 Q59852
 Q59853
 Q59854
 Q59855
 Q59856
 Q59857
 Q59858
 Q59859
 Q59860
 Q59861
 Q59862
 Q59863
 Q59864
 Q59865
 Q59866
 Q59867
 Q59868
 Q59869
 Q59870
 Q59871
 Q59872
 Q59873

Query Match 57.3%; Score 619; DB 9; Length 126;
Best Local Similarity 95.2%; Pred. No. 2.2e-48;
Matches 110; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

1	MSRLDKSVINSALELLNEVGIEGUTTRKLAQKLGVSQPTLYWHVHNKRALLDALAIEML	60
	1	MSRLDKSVINSALELLNEVGIEGUTTRKLAQKLGVSQPTLYWHVHNKRALLDALAIEML
	61	DRHHTHPCPLEGSGWDFLNRKAKSFRCALLSHRDGAKVHLGTRPEKQYETLENOALFL
	61	DRHHTHPCPLEGSGWDFLNRKAKSFRCALLSHRDGAKVHLGTRPEKQYETLENOALFL

RESULT	2	PRT:	259 AA.
Q53901		PRELIMINARY;	
AC	Q53901		
IC	Q53901		
DT	01-NOV-1996	(TREMBREL.	01, CREATED)
DT	01-NOV-1996	(TREMBREL.	01, LAST SEQUENCE UPDATE)
DT	01-NOV-1998	(TREMBREL.	08, LAST ANNOTATION UPDATE)
DE	OR1-14	(ACTIV.	

GN ACT11.
OS STREPTOMYCES COELICOLOR
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; STREPTOMYCETES;
OC STREPTOMYCETACEAE; STREPTOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91347376.
RA FERNANDEZ-MORENO M.A., CABALLERO J.L., HOPWOOD D.A., MALPARTIDA F.;
RT "The act cluster contains regulatory and antibiotic export genes,
RT direct targets for translational control by the bldA trna gene of
RT Streptomyces";
RL CELL 66:769-780(1991).
DR EMBL; M64683; G455360; -.
DR PFAM; PF00440; tetr; 1.
SQ SEQUENCE 259 AA; 28268 MW; 16ECA071 CRC32;

Query Match 13.8%; Score 149.5; DB 9; Length 259;
Best Local Similarity 26.0%; Pred. No. 4.5e-06;
Matches 57; Conservative 35; Mismatches 90; Indels 37; Gaps 8;
QY 4 LDKSKVNSALELNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALD-----AI 57
DB 31 LTQDRIVVTALGILDALSMRRLAQELKTHASLYAHVGNRDELDDIVDLVTEV 90
QY 58 EMLDRHHHTFPCLEGESQWDFLRNKAKSFRCALLSHRDGAKVHLGTRPT-EKQYETLENQ 116
DB 91 EVPE-----PEPGR-WAEQVKEMCSLRMFLAHRDLARIAIDRVPLGPNMGVMGERT 142
QY 117 LAFLCQOQFSLENALYSALVGHHTLGCVLDEQHQVAKEE-----RET 160
DB 143 MNLLRSGLHDLAAYGDLTLSTVTAEALQSSRNPTGREGAQGVFADQLHGVLKSL 202
QY 161 PTDSMPPLRQA--IELFDHQGAEPALFGLGLELIICGL 197
DB 203 PAT-SFPNLVHAGPITSLD---SDRRFELGLEIITAGL 237

RESULT 3
ID 052558 PRELIMINARY; PRT; 242 AA.
AC 052558;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
L1 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
L2 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE REPRESSOR.
GN RIFQ.
OS AMYCOTOPSIS MEDITERRANET.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; PSEUDONOCARDIACEAE;
OC AMYCOTOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8699;
RA KIM C.G., YU T.W., FRYHLE C., HANDA S., FLOSS H.G.;
RL J. BIOL. CHEM. 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=8699;
RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., HUTCHINSON C.R.,
RA TAYLOR M., HOFFMANN D., KIM C.G., ZHANG X., HUTCHINSON C.R.,
RA FLOSS H.G.;
RL CHEM. BIOL. 5:0-0(0002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=8699;
RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., HUTCHINSON C.R.,
RA FLOSS H.G.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF040570; G2792330; -.
SQ SEQUENCE 242 AA; 27049 MW; 324EB570 CRC32;

Query Match 12.0%; Score 129.5; DB 9; Length 242;

Best Local Similarity 26.2%; Pred. No. 0.00025;
Matches 48; Conservative 35; Mismatches 71; Indels 29; Gaps 7;
QY 2 SRLDKSVNSALELNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRAL-----LDALAI 57
DB 27 SEDREKIVATAVRVLDAEGDAKFSRLLAELNVTMPSVYVYVANKDLDLDAVAG 86
QY 58 EM-----LDRHHHTFPCLEGESQWDFLRNKAKSFRCALLSHRDGAKV---HLGTRTEKQY 110
DB 87 ETELPSLD-----DGDNRDELRLARAWRRRTWAHPWAIRCYEYLNIGPSSURF 137
QY 111 EYLENOLAFLCQOQFSLENALYSALVGHHTLGCVLDEQ---EHQV-----AKEERETPTT 163
DB 138 --TECAQAVMACSPLFLKDRSAALNVFQYVYGTATESRWLEHLAETGRTAEFAEAVT 195
QY 164 DSM 166
DB 196 GSM 198

RESULT 4
ID 034892 PRELIMINARY; PRT; 191 AA.
AC 034892;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TRANSCRIPTION REGULATOR.
GN YOB.
OS BACILLUS SUBTILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC BACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RA LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH D.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE; 98044033.
RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
RA ENITIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.,
RA FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N., GHIM S.Y.,
RA GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPPI G.,
RA GUY B.J., HAGA K., HATECH J., HARWOOD C.R., HENAUT A., HILBERT H.,
RA HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L., JORIS B.,
RA KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C., KOBAYASHI Y.,
RA KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M., KURITA K., LAPIDUS A.,
RA LARDINOIS S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H.,
RA MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MEILLADO R.P., MIZUNO M.,
RA MOESTL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,
RA OGAWARA A., OUDEGA B., PARK S.H., PARKO V., POHL T.M., PORTETELLE D.,
RA PERWOLLIG S., PRESCOTT A.M., PRESECAN E., PUJIC P., PURNELLE B.,
RA RAPOPORT G., REY M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,
RA ROCHE B., ROSE M., SADAIE Y., SATO T., SCANLAN E., SCHLEICH S.,
RA SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEROR S.J.,
RA SERROR P., SHIN B.S., SOLDI B., SOROKIN A., TACCONE E., TAKAGI T.,
RA TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,
RA TERPSTRA P., TOGNONI A., TOSATO V., UCHIYAMA S., VANDENBOL M.,
RA VANNIER F., VASSAROTTI A., VIARI A., WABUTTA R., WEDLER H.,
RA WEITZENEGGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,
RA YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis".
RL NATURE 390:249-256(1997).
RN [3]
RP SEQUENCE FROM N.A.

Query Match 10.0%; Score 108.5; DB 9; Length 189;
Best Local Similarity 19.6%; Pred. No. 0.014;
Matches 41; Conservative 32; Mismatches 51; Indels 85; Gaps 5;

```

Db 2 RPTNKRILDAAMQLVKKGYRATITKIEAKANVSEATIFRNFKNQGLVEALLSQHSSN 61
QY 63 HHTHFCPLGESWQDFLNRKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQALFLCQ 122
Db 62 RGSILEQTEGDKYKDL-----HIGT----- 83
QY 123 QGFSLENALVALSAVGHFTLGCVCLEDOEHO-----VAKEERETPTTDSM 166
Db 83 -----CLLELEHRKDKIKISFREPAFMQDVINHVTYEP-QSM 119
QY 167 PPLLRQAIELFDHQ-----AEPALFL 188
Db 120 KOLLVDYLTGMGKGYIQTGNEAHDV 148

RESULT 7
ID O67927 PRELIMINARY; PRT; 192 AA.
AC O67927;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
GN ACRR2.
OS AQUIFEX AEOLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RL aeolicus";
RL NATURE 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000776; G2984362;
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
SQ SEQUENCE 192 AA; 22259 MW; D2E83A94 CRC32;

```

```

Query Match 9.4%; Score 101.5; DB 9; Length 192;
Best Local Similarity 22.7%; Pred. No. 0.06;
Matches 45; Conservative 34; Mismatches 68; Indels 51; Gaps 8;

QY 6 KSKVINSALELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNK-----RALLDALAIEMLD 61
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
  15 KEILSALKLFSKKGFTTINDIAKEVIGTEGAIYRHTFSKEELIKSLLESITKEL-- 73

QY 62 RHHTFCPLGESWQDFLR-----NNAKSFRCALLSHRDGAKVHLGTRTEKQY 110
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
  73 RHKLEVALQGTDEBILESIVDTLIDYAFSPFRLNLYH-----LLKEY 120

QY 111 ETLENO-----LAFICQGFLENALY-----ALSAVGHFTLGCVCLEDOEHOVAKER----- 159
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
  121 GEVKNLPGLIILKFL-----NGLYLRKRLKTYPEITALAVTGSVERVIFKERNFELD 172

QY 159 --ETPTDSMPPLRQAI 174
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
  173 YDEETIKKELKVKLSAI 190

RESULT 8
O53789 PRELIMINARY; PRT; 196 AA.
ID O53789

```

```

AC O53789;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PUTATIVE TRANSCRIPTIONAL REGULATOR.
GN MV040.09.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; MYCOBACTERIA; MYCOBACTERIACEAE;
OC MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA HAMLIN N., CHURCHER C.M.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE; 96181548.
RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37Rv, and comparison with Mycobacterium
RL leprae.";
RL PROG. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL; AL021943; E1253107;
SQ SEQUENCE 196 AA; 21262 MW; 17D86001 CRC32;

Query Match 9.1%; Score 98; DB 9; Length 196;
Best Local Similarity 27.3%; Pred. No. 0.13;
Matches 42; Conservative 24; Mismatches 74; Indels 14; Gaps 5;

QY 2 SRLDKSKVINSALELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDALAIEMLD 61
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
  5 AKLSRESIVEGALFLDREGWDSLTINALATQLTGKPSLYNHVDSLEDLRRVIRVID 64

QY 62 RHHTFCPL-EGESWQDFLRNKAQKSFRCALLSHRDG-----AKVHLGTRTEKQYET--- 113
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
  65 DIITMLNRVGAGRADDVAVLMAGAYR-STAHHPGYSAFTRMPLGGDDPEYTAATRGA 123

QY 113 LENQALFLCQGFLENALYAL----SAVGHFTL 142
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
  124 AAPVIAVLSSYGLDGEQAFYAALFEWSALHGFLV 157

RESULT 9
O51597 PRELIMINARY; PRT; 186 AA.
ID O51597
AC Q51597;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CAM REPRESSOR.
GN CAMR
OS PSEUDOMONAS PUTIDA.
OG PLASMID CAM.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;
OC PSEUDOMONAS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PPG1;
RX MEDLINE; 8623770.
RA KOGA H., ARAMAKI H., YAMAGUCHI E., TAKEUCHI K., HORIUCHI T.,
RA GUNSAIUS I.C.;
RT "camR, a negative regulator locus of the cytochrome P-450cam
RT hydroxylase operon.";
RL J. BACTERIOL. 166:1089-1095(1986).

```

RN SEQUENCE FROM N.A.
 RC STRAIN-PPG1;
 RX MEDLINE: 93326643.
 RA ARAMAKI H., KOGA H., SAGARA Y., HOSOI M., HORIUCHI T.;
 RT "Complete nucleotide sequence of the 5'-exo-hydroxycamphor
 dehydrogenase gene on the CAM plasmid of *Pseudomonas putida* (ATCC
 17453).";
 RL BIOCHIM. BIOPHYS. ACTA 1174:91-94(1993).
 RN [2]
 SEQUENCE FROM N.A.
 RC STRAIN-PPG1;
 RX MEDLINE: 94304997.
 RA ARAMAKI H., SAGARA Y., TAKEUCHI K., KOGA H., HORIUCHI T.;
 RT "Nucleotide sequence of the gene encoding a repressor for the
 cytochrome P-450cam hydroxylase operon on the *Pseudomonas putida* CAM
 plasmid";
 RL BIOCHIMIE 76:63-70(1994).
 DR EMBL: D14680; G473745; -.
 DR PFAM: PF00440; tetr; 1.
 KW PLASMID.
 SQ SEQUENCE 186 AA; 20438 MW; 56E43196 CRC32;

Query Match 8.8%; Score 94.5; DB 9; Length 186;
 Best Local Similarity 28.6%; Pred. No. 0.24;
 Matches 24; Conservative 20; Mismatches 31; Indels 9; Gaps 3;
 QY 6 KSKVNSALELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAIEMLD 61
 DB 4 KOSLLHAARLLSAGKRGDGTMRPICAEGVTPPTLYHHYDGLGHLKAAIDETVQVAE 63
 QY 62 RHH---THFCPLEG--ESQWDLR 80
 DB 64 AYHGTEERGLKGRDGTWATFLQ 87

RESULT 10
 P73186 PRELIMINARY; PRT; 236 AA.
 AC P73186;
 DT 01-FEB-1997 (TREMREL. 02, CREATED)
 DT 01-FEB-1997 (TREMREL. 02, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 26.4 KD PROTEIN.
 OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
 OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
 RN [1]
 RN SEQUENCE FROM N.A.
 RA TABATA S.;
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RN SEQUENCE FROM N.A.
 RX KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
 RA SHIMO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
 RA TABATA S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA RES. 3:109-136(1996).
 DR EMBL: D90904; G1652289; -.
 DR PROSITE: PS01081; HTH_TETR_FAMILY; 1.
 DR PFAM: PF00440; tetr; 1.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 236 AA; 26384 MW; A8D84E37 CRC32;

Query Match 8.8%; Score 94.5; DB 9; Length 236;
 Best Local Similarity 22.5%; Pred. No. 0.33;
 Matches 36; Conservative 24; Mismatches 55; Indels 45; Gaps 5;

QY 8 KVNSALELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAIEMLD 67
 DB 46 RILKALKLFGTKGEGTTTKDLAANVAEGTLFRYFNKKAILEVAT----- 96
 QY 68 CPLEGESWQDFLRNKAISFRCALLSHRQGAHVHGLTRPTEK-OYETLENQALFLCQOQFS 126
 DB 96 -----AGWVEILTD-----LLTELSEMGSKAIAQVMK---RRMFH 128
 QY 127 LENALYALSAVGHFTLGCVCLEDEHQVAKEREETPTTDSM 166
 DB 129 LRENKYLQV-----CFVEAQYHPELREKIQSEIIDKM 161
 RESULT 11
 O67157 PRELIMINARY; PRT; 179 AA.
 ID O67157;
 AC O67157;
 DT 01-AUG-1998 (TREMREL. 07, CREATED)
 DT 01-AUG-1998 (TREMREL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
 DE TRANSCRIPTIONAL REGULATOR (TETR/ACRR FAMILY).
 GN ACRR1.
 OS AQUIFEX AEOLICUS.
 OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-VF5;
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AE000721; G2983549; -.
 SQ SEQUENCE 179 AA; 21554 MW; E359173B CRC32;

Query Match 8.7%; Score 94; DB 9; Length 179;
 Best Local Similarity 21.4%; Pred. No. 0.26;
 Matches 47; Conservative 33; Mismatches 68; Indels 72; Gaps 9;
 QY 6 KSKVNSALELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAIEMLD 65
 DB 4 KERILEVSKELFFKEGKYGQTSVEIVKRNLSKGAFFHFKEKEELI---TEIERTHK 59
 QY 66 HFCPLEGESWQDFLRNKAISFRCALLSHRQGAHVHGLTRPTEKQYETL--ENQALF---- 120
 DB 60 KIISL-----FEENKEKTPPELL-----EMFLEVLYREKKVYIYFLF 96
 QY 120 --LCOQGF-----SLENALYAL-----SAVGHFTLGCVCLEDEHQVAKERE 159
 DB 97 DLLCEKFRNTYFEKIEDAKRRFEKFEKHPPSKAEILSEILGFLQRLILHYVIKERE 156
 QY 160 TPTTDSMPPLRQATLFDHQAEPAPLFLGLELIICGLEK 199
 DB 157 LP---FLAEKLE-----GLKLIFEGVKK 177
 RESULT 12
 O87854 PRELIMINARY; PRT; 217 AA.
 ID O87854;
 AC O87854;
 DT 01-NOV-1998 (TREMREL. 08, CREATED)
 DT 01-NOV-1998 (TREMREL. 08, LAST SEQUENCE UPDATE)

QY	8	KVNSALELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDA	---	LATEMLDRHH	64
		::: : :	: : : :	: :	
Db	14	RIUGAAAEILALRGYSSTSTRDIAAAGVGEQPATYKHFSAKRDILAAALVRLAVEW	----		69

QY 65 THECPLE 71
 Db 69 ----PLE 71

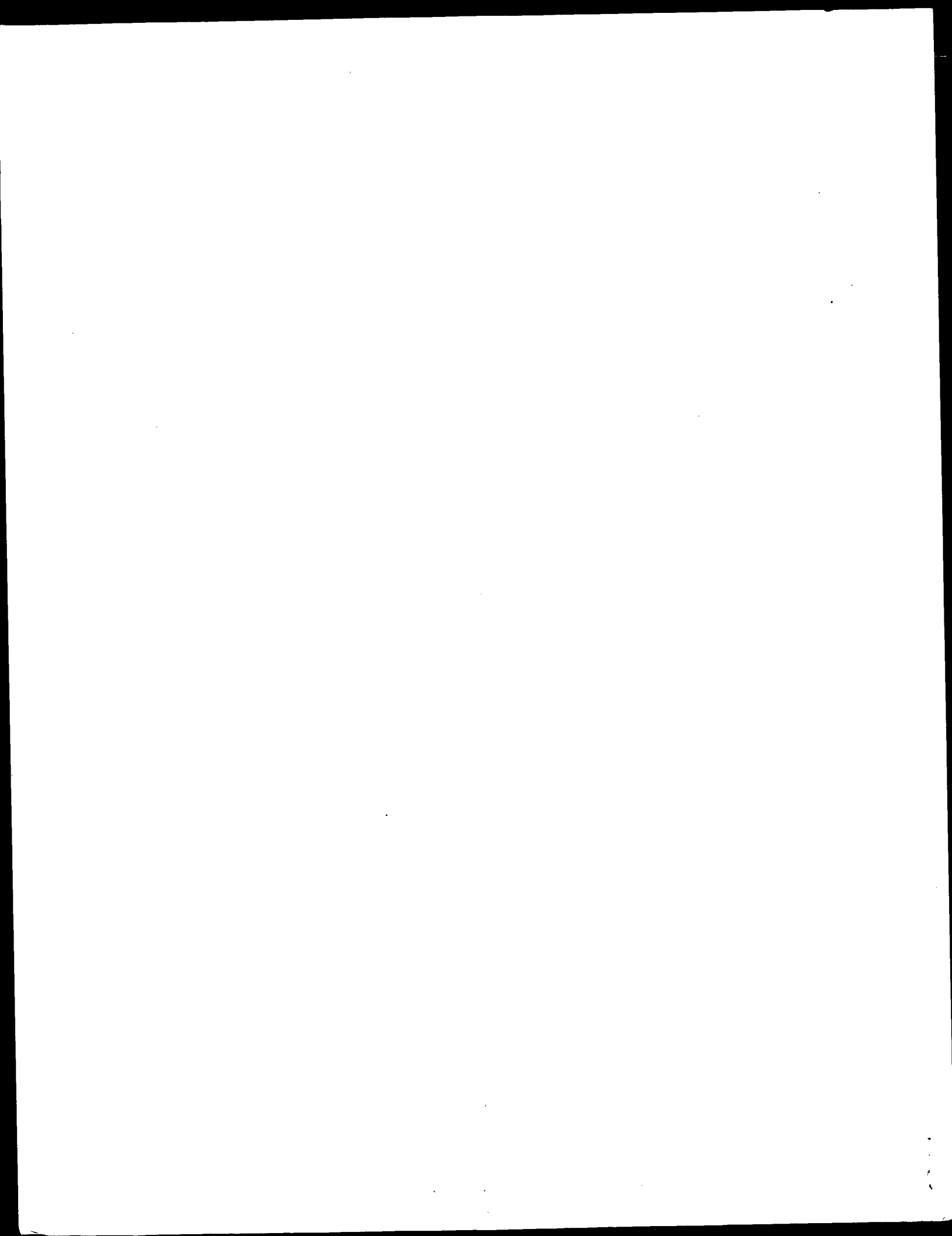
8 KVNSALELLENVIGIEGLTTRKLAQKGVQPTLYWHVKNKRALDLALAIEMLDHRHT 65
 11 KILQAAIEVISEKGLDKASISDIVKKGAGTQGTFFLYFSSKNALIPATAENLLT--HT 66

RESULT 15

O07001
 AC O07001 PRELIMINARY; PRT; 194 AA.
 DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
 DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 22.2 KD PROTEIN.
 GN YVDT.
 OS BACILLUS SUBTILIS.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 CC BACILLUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA DENIZOT F.C.;
 PL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE; 98044033.
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.,
 RA FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N., GHIM S.Y.,
 RA GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPPI G.,
 RA GUY B.J., HAGA K., HALECH J., HARWOOD C.R., HENAUT A., HILBERT H.,
 RA HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L., JORIS B.,
 RA KARAYATA D., KASHARA Y., KLAERR-BLANCHARD M., KLEIN C., KOBAYASHI Y.,
 RA KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M., KURITA K., LAPIDUS A.,
 RA LARDINOIS S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H.,
 RA MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MELLADO R.P., MIZUNO M.,
 RA MOESTL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,
 RA OGIMWARA A., OUDEGA B., PARK S.H., PARRO V., POHL T.N., PORTETELLE D.,
 RA PORWOLLIK S., PRESCOTT A.M., PRESECAN E., PUJIC P., PURNELLE B.,
 RA ROCHE B., ROSE M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,
 RA SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEROR S.J.,
 RA SERROR P., SHIN B.S., SOLDI B., SOROKIN A., TACCONI E., TAKAGI T.,
 RA TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,
 RA TERPSTRA P., TOGNONI A., TOSATO V., UCHIYAMA S., VANDENBOL M.,
 RA VANNIER F., VASSAROTTI A., VIARI A., WAMBUUT R., WEDLER E., WEDLER H.,
 RA WEITZENEGGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,
 RA YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,
 RA YOSHIKAWA H., DANCHIN A.;
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 subtilis."
 RL NATURE 390:249-256(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; 294043; E313126;
 DR EMBL; 295121; E1186136;
 DR PFAM; PF00440; tetr; 1.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 194 AA; 22227 MW; D3457AEC CRC32;

Query Match 8.2%; Score 89; DB 9; Length 194;
 Best Local Similarity 31.0%; Pred. No. 0.8;
 Matches 18; Conservative 17; Mismatches 21; Indels 2; Gaps 1;

Search completed: June 9, 1999, 13:00:33
 Job time: 9152 sec



GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 12:59:17 ; Search time 23.43 Seconds
(without alignments)
178.687 Million cell updates/sec

Title: US-08-486-814-19
 perfect score: 1080
 sequence: 1 MSRLDKSVINSALELLNEV.....FGLELICGLEKQKCESG 207

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database : A_Geneseq_34:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1080	100.0	335	1	R89761	Mutated Tn-10 derl1
2	1080	100.0	207	1	R89764	Mutated Tn-10 derl1
3	1080	100.0	335	1	W08293	Mutated reverse Te
4	1080	100.0	207	1	W08297	Mutated Tet repres
5	1080	100.0	335	1	W13323	Protein encoded by
6	1080	100.0	207	1	W13327	Protein encoded by
7	1058	98.0	335	1	R64808	tTA transactivator
8	1058	98.0	207	1	R89763	Wild-type Tn-10 de
9	1058	98.0	207	1	W08296	Wild-type E.coli T
10	1058	98.0	651	1	W48357	Multi-chimeric tra
11	1058	98.0	651	1	W47583	Multi-chimeric tra
12	1058	98.0	207	1	W13326	Protein encoded by
13	1053	97.5	207	1	R48630	Sequence of the te
14	1053	97.5	297	1	R64809	tTAS transactivato
15	1053	97.5	297	1	R85324	tTAS transactivato
16	1053	97.5	1088	1	R86636	Plasmid pASK75 ope
17	1053	97.5	354	1	W121994	tetracycline trans
18	1053	97.5	349	1	W54311	Pleckstrin homolog
19	1051	97.3	335	1	R85323	tTA transactivator
20	1048	97.0	336	1	W08474	pTet-tTAK encoded
21	1046	96.9	207	1	W08323	Mutated Tet repres
22	1045.5	96.8	206	1	W12401	Mutant tetracyclin
23	1045	96.8	207	1	W08317	Mutated Tet repres
24	1045	96.8	207	1	W08318	Mutated Tet repres
25	1045	96.8	207	1	W08300	Mutated Tet repres
26	1044	96.7	207	1	W08320	Mutated Tet repres
27	1044	96.7	207	1	W08304	Mutated Tet repres
28	1043	96.6	207	1	W08306	Mutated Tet repres
29	1041	96.4	207	1	W08305	Mutated Tet repres
30	1039	96.2	207	1	W08325	Mutated Tet repres
31	1039	96.2	207	1	W08326	Mutated Tet repres
32	1037	96.0	207	1	W08301	Mutated Tet repres
33	1036	95.9	207	1	W08321	Mutated Tet repres
34	1036	95.9	207	1	W08305	Mutated Tet repres
35	1035	95.8	207	1	W08303	Mutated Tet repres
36	1035	95.8	207	1	W08310	Mutated Tet repres
37	1035	95.8	207	1	W08313	Mutated Tet repres
38	1034	95.7	207	1	W08316	Mutated Tet repres
39	1034	95.7	207	1	W08319	Mutated Tet repres
40	1034	95.7	207	1	W08302	Mutated Tet repres
41	1034	95.7	207	1	W08308	Mutated Tet repres
42	1033	95.6	207	1	W08322	Mutated Tet repres
43	1033	95.6	207	1	W08324	Mutated Tet repres

Mutated Tet repres
Mutated Tet repres

ALIGNMENTS

RESULT 1

R89761
ID R89761 standard; Protein; 335 AA.
AC R89761; (first entry)
02-JUL-1996
DE Mutated Tn-10 derived Tet repressor/HSV fusion protein.
DE Tn-10 derived Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
KW fusion protein; gene expression; regulation; inhibition; activation;
KW transcription.
OS Transposon Tn-10/Herpes Simplex Virus.
PN WO9601313-Al.
PD 18-JAN-1996.
PE 29-JUN-1995; U08179.
PF 01-JUL-1994; US-270637.
PR 15-JUL-1994; US-275876.
PR 03-FEB-1995; US-383754.
PR 07-JUN-1995; US-486814.
PA (BUJA/) BUJAR H.
PA (GOSS/) GOSSSEN M.
PI Bujard H, Gossen M;
PI WPI: 96-087666/09.
DR N-PSDB: T11349.
DR N-PSDB: T11349.

New tetracycline-regulated transcription modulators - comprising fusion proteins which bind to tet operator sequences to activate or inhibit transcription
Claim 15; Page 70-71; 112pp; English.
Fusion proteins comprising a first polypeptide which binds to a tet operator sequence in the presence of tetracycline or a tetracycline analogue, operatively linked to a second polypeptide which either activates or inhibits transcription in eukaryotic cells may be used to activate or inhibit transcription. Such proteins may be used to regulate gene expression in cells and may be particularly useful for gene therapy and for expression of gene products in transgenic organisms. Induction of gene expression is rapid, efficient and strong, typically 1000-2000 fold. The inducing agent does not cause pleiotropic effects or cytotoxicity in eukaryotic cells. This sequence is a mutated λ -10 tet repressor operatively linked to a herpes simplex virus virion protein 16 (VP16) activation domain. Sequence 335 AA.

Query Match 100.0%; Score 1080; DB 1: Length 335;
Best Local Similarity 100.0%; Pred. No. 6.8e-115;
Matches 207: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

yy	1	MSRLDSKSVINSALELLNEVGIEGLTRTKLAQKLGVEQPPLYVHVHVKNRALLDALAIEML	60
bb	1	MSRLDSKSVINSALELLNEVGIEGLTRTKLAQKLGVEQPPLYVHVHVKNRALLDALAIEML	60
yy	61	DRHTTHFCPLKGESWQDFLRNKAKSFRCALLSHRNGAKVHSDTRPTEKOYETLENOLAFL	120
bb	61	DRHTTHFCPLKGESWQDFLRNKAKSFRCALLSHRNGAKVHSDTRPTEKOYETLENOLAFL	120
yy	121	CQGFSEIENALVALSAVGHFTLCGVLEDOEHQVAKESERPTTDSMPPLLROAIELFDHQ	180
bb	121	CQGFSEIENALVALSAVGHFTLCGVLEDOEHQVAKESERPTTDSMPPLLROAIELFDHQ	180
yy	181	GASPAFLFGLIELICGLEKQLKCESGS	207
bb	181	GASPAFLFGLIELICGLEKQLKCESGS	207

RESULT	2
RR89764	
ID	R89764 standard; Protein; 207 AA.
AC	R89764;

DT 07-JUL-1996 (first entry)
 DE Mutated Tn-10 derived Tet repressor.
 KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
 KW fusion protein; gene expression; regulation; inhibition; activation;
 KW transcription.
 OS Transposon Tn-10.
 PN W09601313-A1.
 PD 18-JAN-1996.
 PF 29-JUN-1995; U08179.
 PR 01-JUL-1994; US-270637.
 PR 15-JUL-1994; US-275876.
 PR 03-FEB-1995; US-383754.
 PR 07-JUN-1995; US-486814.
 PA (BUJA/) BUJARD H.
 PA (GOS/) GOSSEN M.
 PI Bujard H, Gossen M;
 DR WPI: 96-087666/09.
 DR N-PSDB; T11352.
 PT New tetracycline-regulated transcription modulators - comprising
 PT fusion proteins which bind to tet operator sequences to activate or
 PT inhibit transcription
 PS Claim 44; Page 80; 112pp; English.
 CC Fusion proteins comprising a first polypeptide which binds to a tet
 CC operator sequence in the presence of tetracycline or a tetracycline
 CC analogue, operatively linked to a second polypeptide which either
 CC activates or inhibits transcription in eukaryotic cells may be used
 CC to activate or inhibit transcription. Such proteins may be used to
 CC regulate gene expression in cells and may be particularly useful for
 CC gene therapy and for expression of gene products in transgenic
 CC organisms. Induction of gene expression is rapid, efficient and
 CC strong, typically 1000-2000 fold. The inducing agent does not cause
 CC pleiotropic effects or cytotoxicity in eukaryotic cells. This
 CC sequence is a mutated Tn-10 tet repressor used in a transcription
 CC inhibiting fusion protein.
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred. No. 3.4e-115;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60
 DB 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60
 QY 61 DRHHTFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120
 QY 121 CQGFSLLENALYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIEFLDHQ 180
 DB 121 CQGFSLLENALYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIEFLDHQ 180
 QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 3
 W08293
 ID W08293 standard; Protein; 335 AA.
 AC W08293;
 DE Mutated reverse Tet repressor/VP16 fusion protein.
 KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;
 KW repressor; gene expression; therapy; transgenic animal; disease model.
 OS Synthetic.
 PA Key
 PI Location/Qualifiers
 DR domain 1..207
 FT /note= "reverse tet repressor"
 FT misc_difference 71
 FT /label= substitution
 FT /note= "Lys replaces wild-type Glu"

FT misc_difference 95
 FT /label= substitution
 FT /note= "Asn replaces wild-type Asp"
 FT misc_difference 101
 FT /label= substitution
 FT /note= "Ser replaces wild-type Leu"
 FT misc_difference 102
 FT /label= substitution
 FT /note= "Asp replaces wild-type Gly"
 FT domain 208..335
 FT /note= "C-terminal fragment of VP16 protein"
 PN W09640892-A1.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09049.
 PR 07-JUN-1995; US-485971.
 PA (BADI) BASF AG.
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;
 DR WPI: 97-052305/05.
 PT Nucleic acid encoding tetracycline-inducible transcription
 PT regulatory fusion protein - comprising modified tetracycline
 PT repressor able to bind mutant tet operator, fused to transcription
 PT regulator, useful for modulating eukaryotic gene expression
 PS Example 1; Page 77-78; 117pp; English.
 CC W08293 represents a "reverse" Tet repressor (rtetr), i.e. a wild-type
 CC Tet protein mutated so as to bind its target in the presence
 CC rather than the absence of tetracycline (TC), fused to a C-terminal
 CC polypeptide of VP16 (herpes simplex virus virion protein 16), a
 CC transcriptional activator. The fusion protein forms a Tc-controlled
 CC specificational activator (tTA). The main invention of the
 CC class B tet operator sequences tetO-4C and tetO-6C (see T45711 and
 CC T48478). Modified TetR proteins can be fused to any transcription
 CC regulatory polypeptide and used to control transcription of a tetO-4C
 CC or tetO-6C linked gene. Nucleic acid encoding such a fusion protein
 CC may be introduced into a cell and transcription of the protein can be
 CC controlled by altering the concn. of tetracycline (or an analogue) in
 CC the cell, as appropriate. This ability to modulate gene expression in
 CC a predictable way is very useful in gene therapy and for recombinant
 CC protein prodn. in cultured cells or transgenic animals. The
 CC Tc-inducible system is also useful for the prodn. of transgenic animal
 CC models for the study of disease and also for the study of gene
 CC function e.g. during differentiation. The Tc-inducible system allows
 CC rapid activation of gene transcription without cellular toxicity, high
 CC concns. of inducer are not required.
 SQ Sequence 335 AA;

Query Match 100.0%; Score 1080; DB 1; Length 335;
 Best Local Similarity 100.0%; Pred. No. 6.8e-115;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60
 DB 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60
 QY 61 DRHHTFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120
 QY 121 CQGFSLLENALYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIEFLDHQ 180
 DB 121 CQGFSLLENALYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIEFLDHQ 180
 QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207
 RESULT 4
 W08297
 ID W08297 standard; Protein; 207 AA.
 AC W08297;
 DT 19-MAR-1997 (first entry)

DE Mutated Tet repressor.
KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;
KW repressor; gene expression; therapy; transgenic animal; disease model.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_difference 71 /label= substitution
FT /note= "Lys replaces wild-type Glu"
FT misc_difference 95 /label= substitution
FT /note= "Asn replaces wild-type Asp"
FT misc_difference 101 /label= substitution
FT /note= "Ser replaces wild-type Leu"
FT misc_difference 102 /label= substitution
FT /note= "Asp replaces wild-type Gly"
PN W09640892-A1.
PD 19-DEC-1996.
PR 06-JUN-1996; U09049.
PR 07-JUN-1995; US-485971.
PA (BADI) BASF AG.
PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;
DR WPI; 97-052305/05.
PT Nucleic acid encoding tetracycline-inducible transcription
PT regulatory fusion protein - comprising modified tetracycline
PT repressor able to bind mutant tet operator, fused to transcription
PT regulator, useful for modulating eukaryotic gene expression
PS Example 5; Page 86-87; 117pp; English.
CC W08297 represents a mutated Tet repressor (TetR). This sequence was
CC used to create a fusion protein, comprising the mutated TetR fused to a
CC v-erbA silencer domain (see T45728, W08299). The fusion protein forms
CC a Tc-controlled transcriptional activator (TRA). The main invention of
CC the specification concerns modified TetR proteins that bind to modified
CC class B tet operator sequences tetO-4C and tetO-6C (see T45711 and
CC T48478). Modified tetR proteins can be fused to any transcription
CC regulatory polypeptide and used to control transcription of a tetO-4C
CC or tetO-6C linked gene. Nucleic acid encoding such a fusion protein
CC may be introduced into a cell and transcription of the protein can be
CC controlled by altering the concn. of tetracycline (or an analogue) in
CC the cell, as appropriate. This ability to modulate gene expression in
CC a predictable way is very useful in gene therapy and for recombinant
CC protein prodn. in cultured cells or transgenic animals. The
CC Tc-inducible system is also useful for the prodn. of transgenic animal
CC models for the study of disease and also for the study of gene
CC function e.g. during differentiation. The Tc-inducible system allows
CC rapid activation of gene transcription without cellular toxicity, high
CC concns. of inducer are not required.
SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 3.4e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
DB 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
QY 61 DRHHTHFCPLKGSWQDFLRNKAQKSRFCALLSHRNKGVHSDTRPTKQYETLENQALFL 120
DB 61 DRHHTHFCPLKGSWQDFLRNKAQKSRFCALLSHRNKGVHSDTRPTKQYETLENQALFL 120
QY 121 CQGFSLLENALYALSVAHGFTTLCVLEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180
DB 121 CQGFSLLENALYALSVAHGFTTLCVLEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180
QY 181 GAEPALFGLLELIICGLEKOLKCESGS 207
DB 181 GAEPALFGLLELIICGLEKOLKCESGS 207
RESULT 5
W71327
ID W71327 standard; Protein; 207 AA.
AC W71327;
DT 04-DEC-1998 (first entry)
DE Protein encoded by mutant Tn10-derived tet repressor.

W71323
ID W71323 standard; Protein; 335 AA.
AC W71323;
DT 04-DEC-1998 (first entry)
DE Protein encoded by "reverse" Tet repressor sequence.
KW Tet repressor; rTetR; tetracycline; regulation; expression;
KW Tet operator-linked gene.
OS Synthetic.
PN US5814618-A.
PD 29-SEP-1998.
PF 07-JUN-1995; 485978.
PR 07-JUN-1995; US-485978.
PR 14-JUN-1993; US-076327.
PR 14-JUN-1993; US-076726.
PR 14-JUN-1994; US-260452.
PR 01-JUL-1994; US-270637.
PR 15-JUL-1994; US-275876.
PR 06-FEB-1995; US-383754.
PA (BADI) BASF AG.
PA (KNOL) KNOLL AG.
PI Bujard H, Gossen M;
DR WPI; 98-541795/46.
DR N-PSDB; V60076.
PT Tetracycline based regulation of gene expression - uses a
PT tetracycline operator sequence joined to a gene of interest, the
PT gene of interest being induced in the presence, but not absence of
PT the antibiotic
PS Example 1; Columns 67-68; 63pp; English.
CC The present sequence is encoded by a "reverse" Tet repressor (rTetR),
CC which binds to its target DNA in the presence rather than the absence
CC of tetracycline. The sequence was generated by chemical mutagenesis.
CC rTetR is used in the course of the invention. The specification
CC describes a method for regulating expression of a Tet (tetracycline)
CC operator-linked gene in a cell of a subject. The method comprises
CC introducing into the cell a nucleic acid encoding a fusion protein
CC which inhibits transcription in eukaryotic cells, the fusion protein
CC comprising a polypeptide which binds to a Tet operator sequence,
CC operatively linked to heterologous second polypeptide which inhibits
CC transcription in eukaryotic cells and modulating the concentration of
CC tetracycline (analogue) in the subject. The method is used for the
CC regulation of gene expression system, using tetracycline (analogues).
CC The system enables a gene coupled to the system to be induced in the
CC presence of Tet and then stopped when Tet is removed.
SQ Sequence 335 AA;

Query Match 100.0%; Score 1080; DB 1; Length 335;
Best Local Similarity 100.0%; Pred. No. 6.8e-115;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
DB 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60
QY 61 DRHHTHFCPLKGSWQDFLRNKAQKSRFCALLSHRNKGVHSDTRPTKQYETLENQALFL 120
DB 61 DRHHTHFCPLKGSWQDFLRNKAQKSRFCALLSHRNKGVHSDTRPTKQYETLENQALFL 120
QY 121 CQGFSLLENALYALSVAHGFTTLCVLEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180
DB 121 CQGFSLLENALYALSVAHGFTTLCVLEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180
QY 181 GAEPALFGLLELIICGLEKOLKCESGS 207
DB 181 GAEPALFGLLELIICGLEKOLKCESGS 207
RESULT 6
W71327
ID W71327 standard; Protein; 207 AA.
AC W71327;
DT 04-DEC-1998 (first entry)
DE Protein encoded by mutant Tn10-derived tet repressor.

KW Tet repressor; tetracycline; regulation; expression;
 KW Tet operator-linked gene; tet operator.
 OS Synthetic.
 PN US5814618-A.
 PD 29-SEP-1998.

PF 07-JUN-1995; 485978.
 PR 07-JUN-1995; US-485978.
 PR 14-JUN-1993; US-076327.
 PR 14-JUN-1993; US-076726.
 PR 14-JUN-1994; US-260452.
 PR 01-JUL-1994; US-270637.
 PR 15-JUL-1994; US-275876.
 PR 06-FEB-1995; US-383754.
 PA (BADI) BASF AG.
 PA (KNOL) KNOLL AG.
 PI Bujard H, Gossen M;
 DR WPI: 98-541795/46.
 DR N-PSDB; V60089.

PT Tetracycline based regulation of gene expression - uses a
 PT tetracycline operator sequence joined to a gene of interest, the
 PT gene of interest being induced in the presence, but not absence of
 PT the antibiotic
 PS Claim 9; Columns 81-82; 63pp; English.
 CC The present sequence is encoded by a mutant Tn10-derived tet repressor.
 CC It is used in the course of the invention. The specification describes a
 CC method for regulating expression of a Tet (tetracycline) operator-linked
 CC gene in a cell of a subject. The method comprises introducing into the
 CC cell a nucleic acid encoding a fusion protein which inhibits
 CC transcription in eukaryotic cells, the fusion protein comprising a
 CC polypeptide which binds to a tet operator sequence, operatively linked
 CC to heterologous second polypeptide which inhibits transcription in
 CC eukaryotic cells and modulating the concentration of a tetracycline
 CC (analogue) in the subject. The method is used for the regulation of
 CC gene expression system, using tetracycline (analogues). The system
 CC enables a gene coupled to the system to be induced in the presence of
 CC Tet and then stopped when Tet is removed.
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;
 Best Local Similarity 100.0%; Pred. No. 3 4e-115;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAIEML 60
 DB 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAIEML 60
 OY 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 OY 121 CQOGFSLENALYALSAGHFTLGCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180
 DB 121 CQOGFSLENALYALSAGHFTLGCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180
 OY 181 GAEPAPFLGLELIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 7

ID R64808 standard; Protein; 335 AA.
 AC R64808; 1995 (first entry)
 DE tTA transactivator.
 KW tTA; transactivator; tetracycline-controllable transactivator;
 KW conditional inactivation; homologous recombination; gene expression;
 KW gene regulation; gene therapy; tetracycline-resistance; tetR;
 KW transgenic animal.
 OS Herpes simplex virus K12, KOS.
 PN W09429442-A.
 PD 22-DEC-1994.

PF 14-JUN-1994; U06734.
 PR 14-JUN-1993; US-076327.
 PA (BADI) BASF AG.
 PI Bujard H, Gossen M, Salfeld JG, Voss JW;
 DR WPI: 95-036472/05.
 DR N-PSDB; Q76264.
 PT Regulatory systems using tetracycline-controllable transactivator
 PT (tTA) - useful for conditional inactivation or modulation of
 PT gene expression in a host cell or animal
 PS Disclosure; Page 46-48; 103pp; English.
 CC A 397 bp MluI/PokI fragment of pMSVP16 coding for the C-terminal 130
 CC amino acids of VP16 of HSV was blunted and inserted in pUHD14-1.
 CC The resulting plasmid, pUHD15-1, encoded a tetR-VP16 fusion protein,
 CC or tetracycline-controllable transactivator (tTA), whose sequence is
 CC given in Q76264, and the encoded protein sequence in R64808.
 SQ Sequence 335 AA;

Query Match 98.0%; Score 1058; DB 1; Length 335;
 Best Local Similarity 98.1%; Pred. No. 2.2e-112;
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAIEML 60
 DB 1 MSRLDKSVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAIEML 60
 OY 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 OY 121 CQOGFSLENALYALSAGHFTLGCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180
 DB 121 CQOGFSLENALYALSAGHFTLGCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180
 OY 181 GAEPAPFLGLELIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 8

ID R89763 standard; Protein; 207 AA.
 AC R89763;
 DT 07-JUL-1996 (first entry)
 DE Wild type Tn-10 derived Tet repressor.
 KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
 KW fusion protein; gene expression; regulation; inhibition; activation;
 KW transcription.
 OS Transposon Tn-10.
 PN W09601313-A1.
 PD 18-JAN-1996.
 PF 29-JUN-1995; U08179.
 PR 01-JUL-1994; US-270637.
 PR 15-JUL-1994; US-275876.
 PR 03-FEB-1995; US-383754.
 PR 07-JUN-1995; US-486814.
 PA (BUTAJ) BUTAJARD H.
 PA (GOSSE) GOSSEN M.
 PI Bujard H, Gossen M;
 DR WPI: 96-087666/09.
 DR N-PSDB; T11351.
 PT New tetracycline-regulated transcription modulators - comprising
 PT fusion proteins which bind to tet operator sequences to activate or
 PT inhibit transcription
 PS Claim 38; Page 78; 112pp; English.
 CC Fusion proteins comprising a first polypeptide which binds to a tet
 CC operator sequence in the presence of tetracycline or a tetracycline
 CC analogue, operatively linked to a second polypeptide which either
 CC activates or inhibits transcription in eukaryotic cells may be used
 CC to activate or inhibit transcription. Such proteins may be used to
 CC regulate gene expression in cells and may be particularly useful for
 CC gene therapy and for expression of gene products in transgenic
 CC organisms. Induction of gene expression is rapid, efficient and

CC strong, typically 1000-2000 fold. The inducing agent does not cause
 CC pleiotropic effects or cytotoxicity in eukaryotic cells. This
 CC sequence is the wild type in-10 tet repressor.
 SQ Sequence 207 AA;

Query Match 98.0%; Score 1058; DB 1; Length 207;
 Best Local Similarity 98.1%; Pred. No. 1.1e-112;
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60

QY 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120

QY 121 COQGFLENALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLRQALIEFDHQ 180
 DB 121 COQGFLENALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLRQALIEFDHQ 180

QY 181 GAEPALFGLLEIICGLERKQKCESGS 207
 DB 181 GAEPALFGLLEIICGLERKQKCESGS 207

RESULT 9
 ID W08296 standard; Protein; 207 AA.
 AC W08296;
 DE 19-MAR-1997 (first entry)
 DE Wild-type E.coli Tn10-derived Tet repressor.
 KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;
 KW repressor; gene expression; therapy; transgenic animal; disease model;
 KW Tn10; transposon 10.
 OS Escherichia coli.
 PN W09640892-AL.
 PD 19-DEC-1996.
 PF 06-JUN-1996; W09049.
 PR 07-JUN-1995; US-485971.
 PA (BADI) BASF AG.
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;
 DR WPI; 97-052305/05
 PT Nucleic acid encoding tetracycline-inducible transcription
 PT regulatory fusion protein - comprising modified tetracycline
 PT repressor able to bind mutant tet operator, fused to transcription
 PT regulator, useful for modulating eukaryotic gene expression
 PS Example 4; Page 84-85; 117pp; English.
 CC W08296 represents wild-type Tet repressor (TetR) protein, derived from
 CC transposon 10 (Tn10) of E. coli. The wild-type sequence is used for
 CC the production of a mutated TetR (e.g. so as to bind its target in the
 CC presence rather than the absence of tetracycline). Mutant TetR proteins
 CC can be fused to a transcriptional activator e.g. VP16 (herpes simplex
 CC virus virion protein 16), to form a Tc-controlled transcriptional
 CC activator (tTA). The main invention of the specification concerns
 CC modified TetR proteins that bind to modified class B tet operator
 CC sequences tetO-4C and tetO-6C (see T45711 and T48478). Modified TetR
 CC proteins can be fused to any transcription regulatory polypeptide and
 CC used to control transcription of a tetO-4C or tetO-6C linked gene.
 CC Nucleic acid encoding such a fusion protein may be introduced into a
 CC cell and transcription of the protein can be controlled by altering the
 CC concn. of tetracycline (or an analogue) in the cell, as appropriate.
 CC This ability to modulate gene expression in a predictable way is very
 CC useful in gene therapy and for recombinant protein prodn. in cultured
 CC cells or transgenic animals. The Tc-inducible system is also useful for
 CC the prodn. of transgenic animal models for the study of disease and
 CC also for the study of gene function e.g. during differentiation. The
 CC Tc-inducible system allows rapid activation e.g. during differentiation.
 CC without cellular toxicity, high concns. of inducer are not required.
 SQ Sequence 207 AA;

Query Match 98.0%; Score 1058; DB 1; Length 207;
 Best Local Similarity 98.1%; Pred. No. 1.1e-112;
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60

QY 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120
 DB 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120

QY 121 COQGFLENALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLRQALIEFDHQ 180
 DB 121 COQGFLENALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLRQALIEFDHQ 180

QY 181 GAEPALFGLLEIICGLERKQKCESGS 207
 DB 181 GAEPALFGLLEIICGLERKQKCESGS 207

RESULT 10
 ID W48357 standard; Protein; 651 AA.
 AC W48357;
 DT 20-JUL-1998 (first entry)
 DE Multi-chimeric transactivating factor rTAER fusion protein.
 KW Multi-chimeric transactivating factor; rTAER; tetR;
 KW tetracycline repressor; HSV; oestrogen receptor; promoter;
 KW packaging cell line; retrovirus; retroviral particle; vector;
 KW gene delivery; gene therapy.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Herpes simplex virus.
 OS Chimeric - Mammalia.
 PN W09805759-AL.
 PD 12-FEB-1998.
 PF 06-AUG-1997; U13846.
 PR 07-AUG-1996; US-694652.
 PA (CITY) CITY OF HOPE.
 PA (REGC) UNIV CALIFORNIA.
 PI Chen S, Friedmann T, Miyanochara A, Prussak CE, Yee J;
 DR WPI; 98-145602/13.
 DR N-PSDB; V17756.
 PT New packaging cell lines for pseudotyped retroviral vectors -
 PT comprises sequences encoding retroviral Gag and Pol polypeptide(s)
 PS Disclosure; Fig 4A-C; 84pp; English.
 CC This fusion protein comprises a multi-chimeric transactivator.
 CC designated rTAER, that is composed of (from the N-terminus to the
 CC C-terminus) the Escherichia coli tetR polypeptide, the
 CC transcriptional activation domain of herpes simplex virus VP16, and
 CC the ligand binding domain of the oestrogen receptor. An inducible
 CC expression system of the invention is composed of rTAER and a
 CC minimal promoter (see V17755) derived from the immediate early gene
 CC of cytomegalovirus linked to 7 tandem copies of the tet operator
 CC (tetO) that is the binding site for tetR, which in turn can be
 CC linked to a nucleotide sequence of interest. The invention relates
 CC to packaging cell lines derived from HeLa, D17, MDCK, BHK or
 CC preferably Cf2Th cells and recombinant retroviral particles
 CC produced by them, particularly pseudotyped retroviral particles.
 CC Retroviral particles are produced by inducibly expressing an
 CC envelope protein by linking an envelope protein-encoding nucleotide
 CC sequence to the inducible expression system. The products can be
 CC used for the inducible expression in cells of polypeptides, e.g.
 CC cytotoxic products or therapeutic agents. The activation of the
 CC inducible expression system requires 2 independent signals, thus
 CC reducing the incidence of undesired transcriptional activation.
 SQ Sequence 651 AA;

Query Match 98.0%; Score 1058; DB 1; Length 651;
 Best Local Similarity 98.1%; Pred. No. 5.6e-112;
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 DB 1 MSRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKLSHRNGAKVHSDTRPTKEQYETLENQALFL 120
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKLSHRNGAKVHSDTRPTKEQYETLENQALFL 120
 QY 121 CQOQFSLNAYALSANGVHTGLGCVLEDDQEHQVAKERETPTTDSMPPLRQALIEFLDHQ 180
 DB 121 CQOQFSLNAYALSANGVHTGLGCVLEDDQEHQVAKERETPTTDSMPPLRQALIEFLDHQ 180
 QY 181 GAEPAPFLGLELIICGLEKQKLCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKLCESGS 207

RESULT 11

W47583
 ID W47583 standard; Protein; 651 AA.
 AC W47583; 1998 (first entry)
 DT 21-JUL-1998
 DE Multi-chimeric transactivator tTAER sequence.
 KW Inducible expression system; modulation; cytotoxin; therapeutic;
 KW tTAER; multi-chimeric; transactivator; tetR; Vp16;
 KW ligand-binding domain.
 OS Escherichia coli.
 OS Herpes simplex virus.
 PN W09805754-A2.
 PF 12-FEB-1998.
 PF 29-JUL-1997; U13221.
 PR 07-AUG-1996; US-693940.
 PA (CITY) CITY OF HOPE.
 PA (REGC) UNIV CALIFORNIA.
 PI Chen ST, Friedmann T, Yee JK;
 PP WPI: 98-145597/13.
 DR N-PSDB; V18690.
 PT New inducible expression systems - comprising multi-chimeric
 PT transactivator, induces transcription from promoter in the absence
 PT of first ligand and presence of second ligand
 PS Example 1; Fig 4; 72pp; English.
 CC The sequence is that of multi-chimeric transactivator tTAER.
 CC This is a fusion protein comprising, from N-terminus to C-terminus,
 CC the E.coli tetr polypeptide, the transcription activation domain
 CC of HSV Vp16 and the ligand-binding domain of an oestrogen receptor
 CC from pHE14 (Kumar et al. 1986). DNA encoding it can be used as part of
 CC an inducible expression system which induces transcription from the
 CC promoter in the absence of the first ligand and presence of the
 CC second. The products can be used for the inducible expression in cells
 CC of polypeptides such as cytotoxic products or therapeutic products.
 CC The activation of the inducible expression system requires 2
 CC independent signals, which reduces the incidence of undesired
 CC transcriptional activation.
 SQ Sequence 651 AA;

Query Match 98.0%; Score 1058; DB 1; Length 651;
 Best Local Similarity 98.1%; Pred. No. 5 6e-112;
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 DB 1 MSRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKLSHRNGAKVHSDTRPTKEQYETLENQALFL 120
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKLSHRNGAKVHSDTRPTKEQYETLENQALFL 120
 QY 121 CQOQFSLNAYALSANGVHTGLGCVLEDDQEHQVAKERETPTTDSMPPLRQALIEFLDHQ 180
 DB 121 CQOQFSLNAYALSANGVHTGLGCVLEDDQEHQVAKERETPTTDSMPPLRQALIEFLDHQ 180

QY 181 GAEPAPFLGLELIICGLEKQKLCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKLCESGS 207

RESULT 12

W71326
 ID W71326 standard; Protein; 207 AA.
 AC W71326;
 DT 04-DEC-1998 (first entry)
 DE Protein encoded by wild type Tn10-derived tet repressor.
 KW Tet repressor; tetracycline; regulation; expression;
 KW Tet operator-linked gene; tet operator.
 OS Unidentified.
 PN US814618-A.
 PD 29-SEP-1998.
 PF 07-JUN-1995; 485978.
 PR 07-JUN-1995; US-485978.
 PR 14-JUN-1993; US-076327.
 PR 14-JUN-1993; US-076726.
 PR 14-JUN-1994; US-260452.
 PR 01-JUL-1994; US-270637.
 PR 15-JUL-1994; US-275876.
 PR 06-FEB-1995; US-383754.
 PA (BADI) BASF AG.
 PA (KNOL) KNOLL AG.
 PI Bujard H, Gossen M;
 DR WPI: 98-541795/46.
 DR N-ESDB; V60088.
 PT Tetracycline based regulation of gene expression - uses a
 PT tetracycline operator sequence joined to a gene of interest, the
 PT gene of interest being induced in the presence, but not absence of
 PT the antibiotic
 PS Claim 4; Columns 77-80; 63pp; English.
 CC The present sequence is encoded by wild type Tn10-derived tet repressor.
 CC It is used in the course of the invention. The specification describes a
 CC method for regulating expression of a Tet (tetracycline) operator-linked
 CC gene in a cell of a subject. The method comprises introducing into the
 CC cell a nucleic acid encoding a fusion protein which inhibits
 CC transcription in eukaryotic cells, the fusion protein comprising a
 CC polypeptide which binds to a Tet operator sequence, operatively linked
 CC to heterologous second polypeptide which inhibits transcription in
 CC eukaryotic cells and modulating the concentration of a tetracycline
 CC (analogue) in the subject. The method is used for the regulation of
 CC gene expression system, using tetracycline (analogues). The system
 CC enables a gene coupled to the system to be induced in the presence of
 CC Tet and then stopped when Tet is removed.
 SQ Sequence 207 AA;

Query Match 98.0%; Score 1058; DB 1; Length 207;
 Best Local Similarity 98.1%; Pred. No. 1.1e-112;
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 DB 1 MSRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKLSHRNGAKVHSDTRPTKEQYETLENQALFL 120
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKLSHRNGAKVHSDTRPTKEQYETLENQALFL 120
 QY 121 CQOQFSLNAYALSANGVHTGLGCVLEDDQEHQVAKERETPTTDSMPPLRQALIEFLDHQ 180
 DB 121 CQOQFSLNAYALSANGVHTGLGCVLEDDQEHQVAKERETPTTDSMPPLRQALIEFLDHQ 180
 QY 181 GAEPAPFLGLELIICGLEKQKLCESGS 207
 DB 181 GAEPAPFLGLELIICGLEKQKLCESGS 207

RESULT 13

R48630

ID R48630 standard; Protein; 207 AA.
 AC R48630;
 DT 19-AUG-1994 (first entry)
 DE Sequence of the tetracycline repressor tetr.
 KW Tetracycline repressor; TetR; Tn10; transposon 10.
 OS Escherichia coli.
 PN W09404672-A.
 PD 03-MAR-1994.
 PF 26-AUG-1993; U08230.
 PR 26-AUG-1992; US-935763.
 PA (DNX-) DNX CORP.
 PI Byrne G;
 DR WPI: 94-083191/10.
 DR N-PSDB: Q56710.
 PT Tetracycline repressor-mediated regulation system - useful for
 controlling gene expression in transgenic animals
 PS Disclosure: Page 41-42; 76pp; English.
 CC The inventors claim a construct which comprises an animal promoter
 CC element having a tetracycline repressor (tetR) operator. The
 CC promoter element may be the pPCK promoter which is tissue specific
 CC being expressed selectively in the liver and becoming active
 CC shortly prior to birth. The tetR sequence in the
 CC construct is 3' to a TATA-box sequence and is inserted into the NheI
 CC site of the pPCK promoter element. The entire sequence of the tetR
 CC repressor is given in Q56710/R48630.
 SQ Sequence 207 AA;

Query Match 97.5%; Score 1053; DB 1; Length 207;
 Best Local Similarity 97.6%; Pred. No. 4e-112;
 Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 QY 61 DRHHTHFCPLKGSWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 DB 61 DRHHTHFCPLKGSWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 QY 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFLDHQ 180
 DB 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFLDHQ 180
 QY 181 GAEPAPFLGLELIIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIIICGLEKQKCESGS 207

RESULT 14
 R64809
 ID R64809 standard; Protein; 297 AA.
 AC R64809;
 DT 17-JUL-1995 (first entry)
 DE tTA; transactivator.
 KW tTA; transactivator; tetracycline-controllable transactivator;
 KW conditional inactivation; homologous recombination; gene expression;
 KW gene regulation; gene therapy; tetracycline-resistance; tetR;
 KW transgenic animal.
 OS Herpes simplex virus K12, KOS.
 PN W09429442-A.
 PD 22-DEC-1994.
 PF 14-JUN-1994; U06734.
 PR 14-JUN-1993; US-076327.
 PA (BADI) BASF AG.
 PI Bujard H, Gossen M, Salfeld JG, Voss JW;
 DR WPI: 95-036472/05.
 DR N-PSDB: Q76265.
 PT Regulatory systems using tetracycline-controllable transactivator
 PT (tTA) - useful for conditional inactivation or modulation of
 PT gene expression in a host cell or animal
 PS Disclosure: Page 50-51; 103pp; English.
 CC A DNA fragment of pMSVP16 coding for the C-terminal 97 amino

CC acids of VP16 of HSV was blunted and inserted in pUHD14-1. The
 CC resulting plasmid, pUHD151-1, encoded a tetR-VP16 fusion protein,
 CC or tetracycline-controllable transactivator (smaller version, tTAS),
 CC whose sequence is given in Q76265.
 SQ Sequence 297 AA;

Query Match 97.5%; Score 1053; DB 1; Length 297;
 Best Local Similarity 97.6%; Pred. No. 6.7e-112;
 Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 QY 61 DRHHTHFCPLKGSWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 DB 61 DRHHTHFCPLKGSWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 QY 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFLDHQ 180
 DB 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFLDHQ 180
 QY 181 GAEPAPFLGLELIIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIIICGLEKQKCESGS 207

RESULT 15
 R85324
 ID R85324 standard; Protein; 297 AA.
 AC R85324;
 DT 13-MAR-1996 (first entry)
 DE tTAS transactivator.
 KW tTAS transactivator; tTAS; tet repressor; tetR; virion protein 16; VP16;
 KW gene expression; tetracycline-responsive promoter;
 KW transgenic animal.
 OS Chimeric Escherichia coli;
 OS Chimeric Herpes simplex virus.
 PN US5464758-A.
 PD 07-NOV-1995.
 PF 14-JUN-1993; 076726.
 PR 14-JUN-1993; US-076726.
 PA (BUJA/) BUJARD H.
 PA (GOSSE/) GOSSEN M.
 PI Bujard H, Gossen M;
 DR WPI: 95-392612/50.
 DR N-PSDB: T06868.
 PT Polynucleotide encoding trans-activator fusion protein contg. tet
 PT repressor - used to control expression of gene regulated by minimal
 PT promoter linked to tet operon, and vectors and cells where gene
 PT expression is regulated by tetracycline
 PS Disclosure; Fig 5A-B; 37pp; English.
 CC A fusion protein (R85324) composed of the Escherichia coli Tn10
 CC tet repressor (tetR) and a 127-amino acid C-terminal portion of the
 CC transcription activating domain of herpes simplex virus virion
 CC protein 16 (VP16) is the product of novel tetracycline-controlled
 CC transactivator tTA (T06868). The transactivator is used to regulate
 CC expression of a heterologous gene operably linked to a minimal promoter
 CC derived from human cytomegalovirus promoter IE (see T06869-70) and at
 CC least 1 tet operator (tetO) sequence. On/off regulation of expression
 CC of the heterologous protein by host eukaryotic cells is provided by
 CC varying the medium tetracycline conc. Transgenic animals producing
 CC a desired protein in their milk can also be produced.
 SQ Sequence 297 AA;

Query Match 97.5%; Score 1053; DB 1; Length 297;
 Best Local Similarity 97.6%; Pred. No. 6.7e-112;
 Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60
 QY 61 DRHHTHFCPLKGSWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 DB 61 DRHHTHFCPLKGSWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120
 QY 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFLDHQ 180
 DB 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFLDHQ 180
 QY 181 GAEPAPFLGLELIIICGLEKQKCESGS 207
 DB 181 GAEPAPFLGLELIIICGLEKQKCESGS 207

```

Db 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
QY 61 DRHHTHECPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKQYETILENOLAF 120
Db 61 DRHHTHECPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKQYETILENOLAF 120
QY 121 CQOGFSLLENALYALSAVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180
Db 121 CQOGFSLLENALYALSAVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180
QY 181 GAEPAPFLGLELIICGLEKOLKCESGS 207
Db 181 GAEPAPFLGLELIICGLEKOLKCESGS 207

```

Search completed: June 9, 1999, 12:59:17
 Job time: 9196 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 12:59:50 ; Search time 19.42 seconds
(without alignments)
399.294 Million cell updates/sec

Title: US-08-486-814-19
Perfect score: 1080
Sequence: 1 MSRLDKSVINSALELLNEV.....FGLELIICGLEKOLKCESGS 207

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database : PIR58:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1053	97.5	207	1	RPECTN	E	repressor tetR - Escherichia coli transposon Tn10
2	665	61.6	218	2	S07359		regulatory protein
3	665	61.6	218	2	S30287		regulatory protein
4	597	55.3	126	2	A26948		tetracycline resist
5	512	47.4	219	1	RPECYS		tetracycline repressor
6	505.5	46.8	216	2	QJ1478		regulatory protein
7	505.5	46.8	216	2	S38655		tetR protein - Pse
8	493.5	45.9	216	1	RPECRL		tet repressor prot
9	493.5	45.9	216	1	R77569		tet repressor - Es
10	168.5	15.6	261	2	H07040		hypothetical prote
11	147.5	13.7	259	2	A40046		TetR repressor hom
12	123	11.4	191	2	A69900		hypothetical prote
13	106.5	9.9	210	2	S42419		probable transcrip
14	105.5	9.8	210	2	S42417		probable transcrip
15	103.5	9.6	192	2	C70487		transcription regu
16	102	9.4	189	2	H70042		transcription regu
17	99	9.2	196	2	B70827		hypothetical prote
18	96.5	8.9	246	2	E70861		hypothetical prote
19	94	8.7	179	2	B70391		transcription regu
20	93.5	8.7	236	2	S75298		hypothetical prote
21	92.5	8.6	236	2	D64855		hypothetical prote
22	91.5	8.5	194	2	D64918		glucuronide repres
23	91.5	8.5	213	2	C70035		transcription regu
24	91.5	8.5	213	2	F70946		probable regulato
25	86	8.0	205	2	S29308		hypothetical prote
26	84.5	7.8	543	2	S58095		hypothetical prote
27	83.5	7.7	963	1	A41919		kinesin heavy chai
28	83.5	7.7	216	2	C70649		hypothetical prote
29	83	7.7	195	2	S10899		regulatory protein
30	83	7.7	198	2	E69779		transcription regu
31	83	7.7	2748	2	S57976		nuclear migration
32	82.5	7.6	299	2	S39744		transcription regu
33	82.5	7.6	132	2	E69280		iron-dependent rep
34	82	7.6	212	2	C65165		tkk protein - Esch
35	82	7.6	365	2	S17885		Tcd37 protein - fr
36	81.5	7.5	2472	2	A35715		fodrin alpha chain
37	81	7.5	207	2	H70001		hypothetical prote
38	80.5	7.5	549	1	S19095		transcription fact
39	80.5	7.5	235	2	A64895		conserved hypothet

RESULT 1
RPECIN
repressor tetR - Escherichia coli transposon Tn10
C:Species: Escherichia coli
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 12-Jun-1998
C:Accession: A03576; S02667
R:Postle, K.; Nguyen, T.T.; Berstrand, K.P.
Nucleic Acids Res 12, 4849-4863, 1984
A:Title: Nucleotide sequence of the repressor gene of the Tn10 tetracycline resistance
A:Reference number: A03576; MUID:84247342
A:Accession: A03576
A:Molecule type: DNA
A:Residues: 1-207 <POS>
A:Cross-references: GB:X00694; NID:g43051; PID:g43052
R:Altschmid, L.; Baumeister, R.; Pfeleiderer, K.; Hillen, W.
EMBO J. 7, 4011-4017, 1988
A:Title: A threonine to alanine exchange at position 40 of tet repressor alters the r
A:Reference number: S02667; MUID:89091153
A:Accession: S02667
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-100 <ALT>
A:Cross-references: GB:J01830; NID:gl54845
C:Comment: This protein contains sequences similar to the DNA recognition regions of
C:Genetics:
A:Gene: tetR
C:Superfamily: tetracycline repressor
C:Keywords: antibiotic resistance; DNA binding; transcription regulation
F:25-46/Region: helix-turn-helix #status predicted

Query Match 97.5%; Score 1053; DB 1; Length 207;
Best Local Similarity 97.6%; Pred. No. 2.2e-91;
Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSRLDKSVINSALELLNEVIGLEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60
Db 1 MSRLDKSVINSALELLNEVIGLEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60
Qy 61 DRHHTFCPLKGSQWDFLRNKAHSFRCALLSHRNGAKVHSDTRTEKQYETLENQIAFL 120
Db 61 DRHHTFCPLKGSQWDFLRNKAHSFRCALLSHRNGAKVHSDTRTEKQYETLENQIAFL 120
Qy 121 CQGFSLNAYALSAVGHFTLGCVLDEQEHQVAKEREPTTDSMPPLLRQAIELEFDHQ 180
Db 121 CQGFSLNAYALSAVGHFTLGCVLDEQEHQVAKEREPTTDSMPPLLRQAIELEFDHQ 180
Qy 181 GAEPALFGLLEIICGLEKOLKCESGS 207
Db 181 GAEPALFGLLEIICGLEKOLKCESGS 207

RESULT 2
S07359
regulatory protein tetR - Escherichia coli plasmid RAI
N:Alternate names: tet repressor class D
C:Species: Escherichia coli
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 12-Jun-1998
C:Accession: S07359
R:Unger, B.; Klock, G.; Hillen, W.
Nucleic Acids Res. 12, 7693-7703, 1984
A:Title: Nucleotide sequence of the repressor gene of the RAI tetracycline resistance

A:Reference number: S07359; MUID:85037938
A:Accession: S07359
A:Molecule type: DNA
A:Residues: 1-218 <UNG>
A:Cross-references: EMBL:X01083; NID:g43053; PID:g43054
C:Genetics:
A:Gene: tetr
A:Genome: plasmid
C:Superfamily: tetracycline repressor
C:Keywords: antibiotic resistance; DNA binding; transcription regulation
F:25-46/Region: helix-turn-helix #status predicted

Query Match 61.6%; Score 665; DB 2; Length 218;
Best Local Similarity 63.7%; Pred. No. 4.4e-55;
Matches 128; Conservative 27; Mismatches 46; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
Db 1 MARLNRESVIDAALLELNETGIDGLTRKLAQKLGIEQPTLYWHVKNKRALLDALAVEIL 60

QY 61 DRHHTFCPLKGSWQDFLRNKAQKSHRNGAKVHSDTRTEKQYETLENQALFL 120
Db 61 ARHHDYSLPAAGESWQDFLRNKAQKSHRNGAKVHSDTRTEKQYETLENQALFL 120

QY 121 COQGSFLENALYALSVAHGHTLTCVLEDOHQVAKERETPTTDSMPPLLRQAIELEFDHQ 180
Db 121 TENGFSLRDGLYALSVAHSHFTLGALEQOEHTAALDRPAAPDENLPLLRQAIELEFDHQ 180

QY 181 GAEPALFGLLEIICGLEKQL 201
Db 181 DGEQAFHLGLESIRGFEVOL 201

RESULT 3
S30287
regulatory protein tetr - Salmonella ordonez plasmid pip173
N:Alternate names: tetracycline repressor
C:Species: Salmonella ordonez
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Jun-1998
R:Allard, J.D.; Gibson, M.L.; Vu, L.H.; Nguyen, T.T.; Bertrand, K.P.
Mol. Gen. Genet. 237, 301-305, 1993
A:Title: Nucleotide sequence of class D tetracycline resistance genes from Salmonella ordonez
A:Reference number: S30287
A:Accession: S30287
A:Molecule type: DNA
A:Residues: 1-218 <ALL>
A:Cross-references: EMBL:X65876; NID:g49073; PID:g49075
C:Genetics:
A:Gene: tetr
A:Genome: plasmid
C:Superfamily: tetracycline repressor
C:Keywords: antibiotic resistance; DNA binding; transcription regulation
F:25-46/Region: helix-turn-helix #status predicted

Query Match 61.6%; Score 665; DB 2; Length 218;
Best Local Similarity 63.7%; Pred. No. 4.4e-55;
Matches 128; Conservative 27; Mismatches 46; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
Db 1 MARLNRESVIDAALLELNETGIDGLTRKLAQKLGIEQPTLYWHVKNKRALLDALAVEIL 60

QY 61 DRHHTFCPLKGSWQDFLRNKAQKSHRNGAKVHSDTRTEKQYETLENQALFL 120
Db 61 ARHHDYSLPAAGESWQDFLRNKAQKSHRNGAKVHSDTRTEKQYETLENQALFL 120

QY 121 COQGSFLENALYALSVAHGHTLTCVLEDOHQVAKERETPTTDSMPPLLRQAIELEFDHQ 180
Db 121 TENGFSLRDGLYALSVAHSHFTLGALEQOEHTAALDRPAAPDENLPLLRQAIELEFDHQ 180

QY 181 GAEPALFGLLEIICGLEKQL 201
Db 181 DGEQAFHLGLESIRGFEVOL 201

QY 181 GAEPALFGLLEIICGLEKQL 201
Db 181 DGEQAFHLGLESIRGFEVOL 201

RESULT 4
A26948
tetracycline resistance protein - Haemophilus parainfluenzae
C:Species: Haemophilus parainfluenzae
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 12-Jun-1998
C:Accession: A26948
R:Heuer, C.; Hickman, R.K.; Curiale, M.S.; Hillen, W.; Levy, S.B.
J. Bacteriol. 169, 990-994, 1987
A:Title: Constitutive expression of tetracycline resistance mediated by a Tn10-like element
A:Reference number: A26948; MUID:87137315
A:Accession: A26948
A:Molecule type: DNA
A:Residues: 1-126 <HEU>
A:Cross-references: GB:M15539; NID:g148988; PID:g148989
C:Genetics:
A:Gene: tetr
C:Superfamily: tetracycline repressor
F:25-46/Region: helix-turn-helix #status predicted

Query Match 55.3%; Score 597; DB 2; Length 126;
Best Local Similarity 92.0%; Pred. No. 4.9e-49;
Matches 115; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
Db 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60

QY 61 DRHHTFCPLKGSWQDFLRNKAQKSHRNGAKVHSDTRTEKQYETLENQALFL 120
Db 61 DRHHTFCPLKGSWQDFLRNKAQKSHRNGAKVHSDTRTEKQYETLENQALFL 120

QY 121 COQGF 125
Db 121 ANKVF 125

RESULT 5
RPECYS
tetracycline repressor - Escherichia coli plasmid pSC101
C:Species: Escherichia coli
C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 12-Jun-1998
C:Accession: A03575
R:Brow, M.A.D.; Pesin, R.; Sutcliffe, J.G.
Mol. Biol. Evol. 2, 1-12, 1985
A:Title: The tetracycline repressor of pSC101.
A:Reference number: A03575; MUID:88216101
A:Accession: A03575
A:Molecule type: DNA
A:Residues: 1-219 <BRO>
A:Cross-references: GB:M36272; NID:g150945; PID:g150946
C:Comment: This protein is the repressor of the tetracycline resistance element; its
C:Genetics:
A:Gene: tetr
A:Genome: plasmid
C:Superfamily: tetracycline repressor
C:Keywords: antibiotic resistance; DNA binding; transcription regulation
F:24-46/Domain: DNA binding #status predicted <DBN>
F:25-46/Region: helix-turn-helix #status predicted

Query Match 47.4%; Score 512; DB 1; Length 219;
Best Local Similarity 51.7%; Pred. No. 8.9e-41;
Matches 105; Conservative 33; Mismatches 61; Indels 4; Gaps 1;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
Db 1 MNKLOREAVIRTALELLNDVGMGLTTRRAERLGVQVQALYWHVKNKRALLDALAEAML 60

QY 61 DRHHTFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120
 Db 61 TINTHTTPRDDDDWRSFLRGNACSFRRALLAYRDGARIHAGTRPAAPQMEKADQAQLRFL 120
 QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKERE---TPTTDSMP--PILRQAIEL 176
 Db 121 CDAGFSAGDATYALMALSYFTVGAVLEEQASEDAERGEDQLTTSASTPARLOSAMKI 180
 QY 177 FDHGAEPAPFLGELIICGLEK 199
 Db 181 VYEGGPDAAFERGLIIGGLEK 203

RESULT 6
 QJ1478
 regulatory protein tetr - Escherichia coli transposon Tn1721
 N;Alternate names: tetracycline resistance repressor
 C;Species: Escherichia coli
 C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 12-Jun-1998
 C;Accession: JQ1478; S02668; S24113
 R;Altmeyer, H.; Cresnar, B.; Greck, M.; Schmitt, R.
 Gene 111, 11-20, 1992
 A;Title: Complete nucleotide sequence of Tn1721: gene organization and a novel gene product
 A;Reference number: JQ1475; MUID:92192465
 A;Accession: JQ1478
 A;Molecule type: DNA
 A;Residues: 1-216 <ALL>
 A;Cross-references: EMBL:X61367; NID:g48194; PID:g48198
 A;Note: the authors translated the codon GAG for residue 19 as Asp and CAG for residue 8
 R;Altschmid, L.; Baumeister, R.; Pfeleiderer, K.; Hillen, W.
 EMBO J. 7, 4011-4017, 1988
 A;Title: A threonine to alanine exchange at position 40 of tet repressor alters the repressor
 A;Reference number: S02667; MUID:89091153
 A;Accession: S02668
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-100 <ALT>
 C;Genetics: tetR
 A;Gene: tetR
 C;Superfamily: tetracycline repressor
 F;25-46/Region: helix-turn-helix #status predicted

Query Match 46.8%; Score 505.5; DB 2; Length 216;
 Best Local Similarity 51.5%; Pred. No. 3.5e-40;
 Matches 104; Conservative 35; Mismatches 60; Indels 3; Gaps 2;

QY 1 MSRLDKSKVINSALLENVEGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAEML 60
 Db 1 MTKLPQNTVIRAAIDLLENEGVVDGLTTRKLAERLGVQOPALYWHFRNKRALLDALAEML 60
 QY 61 DRHHTFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120
 Db 61 AENHTSVPRADDWRSFLRGNACSFRRALLAYRDGARIHAGTRPAAPQMEKADQAQLRFL 120
 QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKEREPTTDSMP--PILRQAIELFD 178
 Db 121 CEAGFSAGDAVNALMTISYFTVGAVLEEQAGSDAGER-GGTVEQAPLSPLLRADDAFD 179
 QY 179 HOGAEPAPFLGELIICGLEK 200
 Db 180 EAGPDAAFEQGLAVIVDGLAKR 201

RESULT 7
 S38655
 tetR protein - Pseudomonas aeruginosa
 C;Species: Pseudomonas aeruginosa
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 12-Jun-1998
 C;Accession: S38655
 R;Trueman, P.; Sharpe, G.S.; Barth, P.T.
 submitted to the EMBL Data Library, November 1993

A;Reference number: S38655
 A;Accession: S38655
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-216 <TRU>
 A;Cross-references: EMBL:X75761; NID:g415984; PID:g415985
 C;Superfamily: tetracycline repressor
 F;25-46/Region: helix-turn-helix #status predicted

Query Match 46.8%; Score 505.5; DB 2; Length 216;
 Best Local Similarity 51.5%; Pred. No. 3.5e-40;
 Matches 104; Conservative 35; Mismatches 60; Indels 3; Gaps 2;

QY 1 MSRLDKSKVINSALLENVEGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAEML 60
 Db 1 MTKLPQNTVIRAAIDLLENEGVVDGLTTRKLAERLGVQOPALYWHFRNKRALLDALAEML 60
 QY 61 DRHHTFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120
 Db 61 AENHTSVPRADDWRSFLRGNACSFRRALLAYRDGARIHAGTRPAAPQMEKADQAQLRFL 120
 QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKEREPTTDSMP--PILRQAIELFD 178
 Db 121 CEAGFSAGDAVNALMTISYFTVGAVLEEQAGSDAGER-GGTVEQAPLSPLLRADDAFD 179
 QY 179 HOGAEPAPFLGELIICGLEK 200
 Db 180 EAGPDAAFEQGLAVIVDGLAKR 201

RESULT 8
 RPECRI
 tet repressor protein (Tn 1721) - Escherichia coli plasmid RP1
 C;Species: Escherichia coli
 C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 12-Jun-1998
 C;Accession: A03574
 R;Waters, S.H.; Rogowsky, P.; Grinsted, J.; Altenbuchner, J.; Schmitt, R.
 Nucleic Acids Res. 11, 6089-6105, 1983
 A;Title: The tetracycline resistance determinants of RP1 and Tn1721: nucleotide sequence
 A;Reference number: A93486; MUID:83299270
 A;Accession: A03574
 A;Molecule type: DNA
 A;Residues: 1-216 <WAT>
 A;Cross-references: GB:X00006; NID:g42508; PID:g42509
 C;Genetics: tetR
 A;Gene: tetR
 A;Genome: plasmid
 C;Superfamily: tetracycline repressor
 C;Keywords: antibiotic resistance; DNA binding; transcription regulation
 F;25-46/Region: helix-turn-helix #status predicted

Query Match 45.9%; Score 495.5; DB 1; Length 216;
 Best Local Similarity 50.5%; Pred. No. 3.1e-39;
 Matches 102; Conservative 36; Mismatches 61; Indels 3; Gaps 2;

QY 1 MSRLDKSKVINSALLENVEGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAEML 60
 Db 1 MTKLPQNTVIRAAIDLLENEGVVDGLTTRKLAERLGVQOPALYWHFRNKRALLDALAEML 60
 QY 61 DRHHTFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120
 Db 61 AENHTSVPRADDWRSFLRGNACSFRRALLAYRDGARIHAGTRPAAPQMEKADQAQLRFL 120
 QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKEREPTTDSMP--PILRQAIELFD 178
 Db 121 CEAGFSAGDAVNALMTISYFTVGAVLEEQAGSDAGER-GGTVEQAPLSPLLRADDAFD 179
 QY 179 HOGAEPAPFLGELIICGLEK 200
 Db 180 EAGPDAAFEQGLAVIVDGLAKR 201

```
RESULT 9
177569
tet repressor - Escherichia coli
C:Species: Escherichia coli
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 12-Jun-1998
C:Accession: I77569
R:Tovar, K.; Ernst, A.; Hillen, W.
Mol. Gen. Genet. 215, 76-80, 1988
A:Title: Identification and nucleotide sequence of the class E tet regulatory elements
A:Reference number: I57766; MUID:89201249
A:Accession: I77569
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: EMBL:X14035; NID:943055; PID:g43056
C:Superfamily: tetracycline repressor
A:45-46/Region: helix-turn-helix #status predicted

Query Match 45.7%; Score 493.5; DB 2; Length 211;
Best Local Similarity 50.2%; Pred. No. 4.6e-39;
Matches 101; Conservative 29; Mismatches 68; Indels 3; Gaps 1;

QY 1 MSRLDKSVNSALELNEVGIEGLTRKLAQKLGVEOPTLYWHVKNRALLDALAIEML 60
D: 1 MARSLDDVISMAFLDSEGLIEGLTRKLAQKLGVEOPTLYWHVKNRQTLMMNLSEAIL 60
QY 61 DRHHTHFCPLKGSQWDFLRNKAFCRALLSHRNGAKVHSDTRPTKEQYETLENQAF 120
D: 61 AKHHTRSAPLPTESWQOFLQENALSFRRKALLVHRDGLHIGTSTPTPPQFEAQAQLRCL 120
QY 121 COQGSFLENALYALSAGVHFTLGCVLDEQEHQVAKEREPTTDSMPPLLRQALIELFDHQ 180
D: 121 CDAGFSVEALFLOSISGHFTLGAVLEE--QATNQIENNHIIDAPPLQEAENIQART 177
QY 181 GAEPALFGLGLICGLKQQL 201
D: 178 SAEMAFHFLKSLIFGSAQL 198

RESULT 10
H70740
Hypothetical protein Rv1353c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998
C:Accession: H70740
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feilwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nature 393, 537-544, 1998
A:Authors: Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, A.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: H70740
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-261 <COL>
A:Cross-references: GB:275555; GB:AL123456; NID:93261608; PID:e250355; PID:g1419056
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1353c

Query Match 15.68; Score 168.5; DB 2; Length 261;
Best Local Similarity 27.08; Pred. No. 1.5e-08;
Matches 61; Conservative 32; Mismatches 78; Indels 55; Gaps 6;

QY 4 LDKSKVNSALELNEVGIEGLTRKLAQKLGVEOPTLYWHVKNRALLDALAIEMLDHR 63
D: 16 INPEDIISGAFELAQOVSIDNLSMPLKGLGVGTSTYIWFRRKDDLLNATDRALSKY 75
QY 64 HTHFCPLKGSQWDFLRNKAFCR--CALL-----SHRNGAKVHSDTRP 105

RESULT 11
A40046
Tet repressor homolog actII-1 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 12-Jun-1998
C:Accession: A40046
R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.
Cell 66, 769-780, 1991
A:Title: The act cluster contains regulatory and antibiotic export genes, direct targ
A:Reference number: A40046; MUID:91347376
A:Accession: A40046
A:Molecule type: DNA
A:Residues: 1-259 <FER>
A:Cross-references: GB:M64683; NID:g153143; PID:g455360
C:Genetics:
A:Start codon: GTG
F:52-73/Region: helix-turn-helix #status predicted

Query Match 13.7%; Score 147.5; DB 2; Length 259;
Best Local Similarity 26.0%; Pred. No. 1.3e-06;
Matches 57; Conservative 35; Mismatches 90; Indels 37; Gaps 8;

QY 4 LDKSKVNSALELNEVGIEGLTRKLAQKLGVEOPTLYWHVKNRALLDAL--AI 57
D: 31 LTQDRIVVTALGILDAEGLDALSRRRLAQELKTHASLYAHVGNRELLDLVFDVILTEV 90
QY 58 EMLDRHHTHFCPLKGSQWDFLRNKAFCRALLSHRNGAKVHSDTRPT-KEQYETLENQ 116
D: 91 EYPE-----PEPGR-WAEQVKEMCRSLRMLFLAHRDAIDRVPLGPNMGVMGERT 142
QY 117 LAFLCQOQGSFLENALYALSAGVHFTLGCVLDEQEHQVAKER-----RET 160
D: 143 MNLSSGGLHDELAAYGGDLSTFTVAELQSSRNPTQEGREQAGVAFDQLHGKYLKSL 202
QY 161 PTTDSMPPLLRQA--IELFDHQAEPALFGLGLIICGL 197
D: 203 PAT-SFPNVLHLAGPITSLD---SDRRFELGLEIIAGL 237

RESULT 12
A69900
Hypothetical protein yobS - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1998
C:Accession: A69900
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Gh
wood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M
Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Laube
C.; Authors: Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maugel, C.; Med
K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Por
Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.; Schl
A:Authors: Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sroror, S.J.; Ser
anakhshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vanden
anahshi, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikaw
A:Authors: Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
```

A:Reference number: A69580; MUID:98044033

A:Accession: A69900

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-191 <KUN>

A:Cross-references: GB:299114; GB:AL009126; NID:g2634230; PID:e1185379; PID:g2634300

A:Experimental source: strain 168

C:Genetics:

A:Gene: yobS

Query Match 11.4%; Score 123; DB 2; Length 191;

Best Local Similarity 24.1%; Pred. No. 0.00018;

Matches 49; Conservative 38; Mismatches 74; Indels 42; Gaps 8;

QY 4 LDKSKVNSALELLNEVGIEGLTRKLAOKLGVEOPTLYWHVKNKRALDALAIAI-----EM 59

DB 7 LTKMIVDAAEATADQGVNGVSLAALSKMNVPPSLYNHNGLOAIRAELAVRGLTKL 66

QY 60 LDRHHTHFCPLKGES-----WDPLRNKAKSFRCALLSHRNKAKVHSDTRPTKQYE 111

DB 67 FDQWADSVTERKGD SAMLSAHAYVDFAIENPGYEAALL-----KVH-----DKRTE 114

QY 112 TLENQALFCCQGSFLENALYALSNGHFTLG--CVLEQEHQVAK--ERETPTTDSMP 167

DB 115 IVSDQIVCLVTK-LLENGYASEKTAIHATRGRLSLHGTVLIAKEAFEREEDILESLS 173

QY 168 PLLRQAJELFDHQAEPAPFLFGL 190

DB 174 FSIR-----TFLSGL 183

RESULT 13

S42419

probable transcription repressor mtrR - Neisseria gonorrhoeae (strain FA19)

C:Species: Neisseria gonorrhoeae

A:Variety: strain FA19

C:Date: 07-May-1998 #sequence_revision 15-May-1998 #text_change 18-Sep-1998

C:Accession: S42419; S40250

R:Pan, W.; Spratt, B.G.

Mol. Microbiol. 11, 769-775, 1994

A:Title: Regulation of the permeability of the gonococcal cell envelope by the mtr system

A:Reference number: S42417

A:Accession: S42419

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-210 <PAN>

A:Cross-references: EMBL:225797; NID:g452332; PID:g438189

A:Experimental source: strain FA19

C:Genetics:

A:Gene: mtrR

C:Superfamily: probable transcription repressor mtrR

C:Keywords: DNA binding; transcription regulation

Query Match 9.9%; Score 106.5; DB 2; Length 210;

Best Local Similarity 26.8%; Pred. No. 0.0069;

Matches 51; Conservative 23; Mismatches 73; Indels 43; Gaps 8;

QY 6 KSKVNSALELLNEVGIEGLTRKLAOKLGVEOPTLYWHVKNKRALDALAIAIEMDRHHT 65

DB 12 KEHLMALAEITYRKGIARTSLNEIAQAAGVTRGALYWHFKNEDLFDALFORICDDIEN 71

QY 66 ----HFCPLKGSWODFLRNKAKSFRCALLSHRNKAKVHSDTRPTKQYETLENOLAFLC 121

DB 72 CIAQDAADAEGGSWTVF-RHTLLHFFERLQSNDIYKHF-----NILFLKC 116

QY 122 QGFSLENALYALSNGHFTLGCVLEQEHQVAKERTPTTDSMPPLLRQAI-----LFD 178

DB 117 EH--TEQNA--AVIAIA-----RKHOAIWREKIT-----AVLTAVERNQDLAD 155

QY 179 HOGAEPAPFLF 188

DB 156 DLDKETAVIF 165

RESULT 14

S42417

probable transcription repressor mtrR - Neisseria gonorrhoeae (strain CH95)

C:Species: Neisseria gonorrhoeae

A:Variety: strain CH95

C:Date: 07-May-1998 #sequence_revision 15-May-1998 #text_change 18-Sep-1998

C:Accession: S42417; S40251

R:Pan, W.; Spratt, B.G.

Mol. Microbiol. 11, 769-775, 1994

A:Title: Regulation of the permeability of the gonococcal cell envelope by the mtr system

A:Reference number: S42417

A:Accession: S42417

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-210 <PAN>

A:Cross-references: EMBL:225796; NID:g438190; PID:g438191

A:Experimental source: CH95

C:Genetics:

A:Gene: mtrR

C:Superfamily: probable transcription repressor mtrR

C:Keywords: DNA binding; transcription regulation

Query Match 9.8%; Score 105.5; DB 2; Length 210;

Best Local Similarity 26.8%; Pred. No. 0.0086;

Matches 51; Conservative 23; Mismatches 73; Indels 43; Gaps 8;

QY 6 KSKVNSALELLNEVGIEGLTRKLAOKLGVEOPTLYWHVKNKRALDALAIAIEMDRHHT 65

DB 12 KEHLMALAEITYRKGIARTSLNEIAQAAGVTRGALYWHFKNEDLFDALFORICDDIEN 71

QY 66 ----HFCPLKGSWODFLRNKAKSFRCALLSHRNKAKVHSDTRPTKQYETLENOLAFLC 121

DB 72 CIAQDAADAEGGSWTVF-RHTLLHFFERLQSNDIYKHF-----NILFLKC 116

QY 122 QGFSLENALYALSNGHFTLGCVLEQEHQVAKERTPTTDSMPPLLRQAI-----LFD 178

DB 117 EH--TEQNA--AVIAIA-----RKHOAIWREKIT-----AVLTAVERNQDLAD 155

QY 179 HOGAEPAPFLF 188

DB 156 DLDKETAVIF 165

RESULT 15

C70487

transcription regulator Tetr/AcrR family - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Jul-1998

C:Accession: C70487

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98198666

A:Accession: C70487

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-192 <AOE>

A:Cross-references: GB:AE000776; NID:g2984355; PID:g2984362; GB:AE000657

A:Experimental source: strain VF5

C:Genetics:

A:Gene: acrR2

Query Match

Best Local Similarity 9.6%; Score 103.5; DB 2; Length 192;

Matches 45; Conservative 35; Mismatches 67; Indels 51; Gaps 8;

QY 6 KSKVNSALELLNEVGIEGLTRKLAOKLGVEOPTLYWHVKNK-----RALLDALAIEMD 61

```

| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
15 K E I L S S A K L F S K G F K E T T I K D I A K E V G I T E G A I Y R H F T S K E E I I K S L L E S I T K E L -- 73
Qy 62 R H H T H F C P L K G E S W Q D F L R ----- N K A K S F R C A L L S H R N G A K V H S D T R P T E K Q Y 110
Db 73 R H K L E V A L Q R G E T D E E I L E S I V D T L I D Y A F S N P E S F R F L N L Y H ----- L L K E Y 120
Qy 111 E T L E N Q ----- L A F L C O O G F S L E N A L Y --- A L S A V G H F T L G C V L E D O E H O V A K E E R ----- 159
Db 121 G E V K N L P C E L I L K E L ----- N G L Y K R K L K T Y P E I A L A V V T G S V E R V F I F K E R N F L D 172
Qy 159 -- E T P T T D S M P P L L R Q A I 174
Db 173 Y D E E T I R K K L K K V L K S A I 190

```

Search completed: June 9, 1999, 12:59:50
Job time: 9169 sec

PO8238 homo sapien
P75342 mycoplasma

44 73 5 6.8 723 1 H59B_HUMAN
45 73 6.8 1244 1 Y307_MYCPN

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 13:01:02 ; Search time 16.52 Seconds
(without alignments)
336.316 Million cell updates/sec

Title: US-08-486-814-19
Perfect score: 1080
Sequence: 1 MSRLDKSKVNSALELLNEV.....FGLELICGLEKOLKCESGS 207

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1053	97.5	207	1	TER2_ECOLI	P04483 escherichia
2	665	61.6	218	1	TER8_PASPI	P51562 pasteurella
3	663	61.4	207	1	TER8_PASPI	P51561 pasteurella
4	660	61.1	217	1	TER4_ECOLI	P09164 escherichia
5	512	47.4	219	1	TER3_ECOLI	P03039 escherichia
6	505.5	46.8	216	1	TER1_ECOLI	P03038 escherichia
7	493.5	45.7	211	1	TER5_ECOLI	P21337 escherichia
8	473	43.8	210	1	TER7_VIBAN	P51560 vibrio angu
9	168.5	15.6	261	1	Y07H_MYCTU	Q11023 mycobacteri
10	106.5	9.9	210	1	MTRR_NEIGO	P39897 neisseria g
11	92.5	8.6	236	1	YCFQ_ECOLI	P75952 escherichia
12	91.5	8.5	196	1	UDR_ECOLI	Q59431 escherichia
13	86	8.0	955	1	PHSL_IPOBA	P27598 ipomea bat
14	84.5	7.8	577	1	THTL_SCHPO	Q09684 schizosacch
15	83.5	7.7	963	1	KINH_HUMAN	P33176 homo sapien
16	83	7.7	195	1	BETI_ECOLI	P17446 escherichia
17	82.5	7.6	299	1	YWFK_BACSU	P39647 bacillus su
18	82	7.6	212	1	TTK_ECOLI	P06969 escherichia
19	81.5	7.5	2472	1	SPCN_HUMAN	Q13813 homo sapien
20	80.5	7.5	549	1	CFIA_DROME	P16241 drosophila
21	80	7.4	188	1	YP23_STAAG	P23217 staphylococ
22	79.5	7.4	1391	1	N157_YEAST	P40064 saccharomyc
23	79	7.3	187	1	Y893_HAEIN	P44923 haemophilus
24	78.5	7.3	770	1	PRTP_VZVD	P09284 varicella-z
25	78.5	7.3	218	1	TTK_HAEIN	P29280 haemophilus
26	78	7.2	215	1	ACRR_ECOLI	P34000 escherichia
27	78	7.2	519	1	N1FL_AZOVI	P30663 azotobacter
28	77.5	7.2	212	1	YCDC_ECOLI	P75899 escherichia
29	77	7.1	192	1	BM3R_BACME	P43506 bacillus me
30	77	7.1	2748	1	NUM1_YEAST	Q00402 saccharomyc
31	77	7.1	673	1	PLZF_HUMAN	Q05516 homo sapien
32	76.5	7.1	838	1	PHSH_SOLTU	P32811 solanum tub
33	75	6.9	1324	1	CUT3_SCHPO	P41004 schizosacch
34	75	6.9	2319	1	FAB_MOUSE	Q06194 mus musculu
35	75	6.9	1509	1	MYSN_ACACA	P05659 acanthamoeb
36	74.5	6.9	557	1	ATRA_ECOLI	P03959 escherichia
37	74	6.9	2133	1	FAB_PIG	P12263 sus scrofa
38	74	6.9	474	1	LMN3_MOUSE	P48680 mus musculu
39	74	6.9	2116	1	MY52_DICDI	P08799 dictyosteli
40	74	6.9	892	1	YJ06_YEAST	P40985 saccharomyc
41	74	6.9	1207	1	YJ08_CAEEL	P34402 caenorhabdi
42	73.5	6.8	323	1	FRA2_CHICK	P18625 gallus gall
43	73.5	6.8	728	1	HS9A_CHICK	P11501 gallus gall

ALIGNMENTS

```

RESULT 1
TER2_ECOLI
ID TER2_ECOLI STANDARD; PRT; 207 AA.
AC P04483;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS B (TRANSPONSON TN10).
GN TETR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84247342.
RA POSTLE K., NGUYEN T.T., BERTRAND K.P.;
RL NUCLEIC ACIDS RES. 12:4849-4863(1984).
RN [2]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE; 84005886.
RA BERTRAND K.P., POSTLE K., WRAY L.V. JR., REZNIKOFF W.S.;
RL GENE 23:149-156(1983).
RN [3]
RP MUTAGENESIS.
RX SMITH L.D., BERTRAND K.P.;(1988).
RL J. MOL. BIOL. 203:949-959(1988).
CC -1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETA) PROMOTER OPERATOR SITES.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; X00694; G43052; -.
DR EMBL; J01830; G154846; -.
DR PIR; A03576; RPECTN.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
DR TRANSCRIPTION REGULATION; REPRESSOR; TRANSPOSABLE ELEMENT;
KW DNA-BINDING; ANTIBIOTIC RESISTANCE.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
FT MUTAGEN 64 64 SIMILARITY).
FT MUTAGEN 82 82 H->Y: REDUCES AFFINITY FOR TETRACYCLINE
FT MUTAGEN 82 82 N->H: REDUCES AFFINITY FOR TETRACYCLINE
FT MUTAGEN 103 103 T->I: REDUCES AFFINITY FOR TETRACYCLINE
FT SEQUENCE 207 AA; 23354 MW; 4AB4DD5A CRC32;

```

```

Query Match 97.5%; Score 1053; DB 1; Length 207;
Best Local Similarity 97.6%; Pred. No. 3 4e-92;
Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Oy 1 MSRLDKSKVNSALELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60
Db 1 MSRLDKSKVNSALELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60
Oy 61 DRHHTHCPLKGSWQDFLRNKAQSFRCALLSHRNGAKVHSIDRPTKQYETLENLAFL 120
Db 61 DRHHTHCPLKGSWQDFLRNKAQSFRCALLSHRNGAKVHSIDRPTKQYETLENLAFL 120

```

```
QY 121 COQGSLENAIALYSAVGHFTLCVLEQDEHOVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
D 121 COQGSLENAIALYSAVGHFTLCVLEQDEHOVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
QY 181 GAEPAPFLGLELIICGLEKQKLCESGS 207
D 181 GAEPAPFLGLELIICGLEKQKLCESGS 207

RESULT 2
TER8_PASPI
ID TER8_PASPI STANDARD; PRT; 218 AA.
AC P51562;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS H.
CN TETR.
OS PASTEURILLA PISCICIDA.
OG PLASMID P59351.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC PASTURELLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RA HANSEN L.M., MCMURRY L.M., LEVY S.B., HIRSH D.C.;
RA ANTIMICROB. AGENTS CHEMOTHER. 37:2699-2705(1993).
CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETA) PROMOTER OPERATOR SITES.
CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; D15172; G575938; -.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;
KW ANTI-BIOTIC RESISTANCE; PLASMID.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT METAL 100 100 SIMILARITY).
FT SEQUENCE 218 AA; 24419 MW; 3C57D4AB CRC32;

Query Match 61.6%; Score 665; DB 1; Length 218;
Best Local Similarity 63.7%; Pred. No. 1.2e-55;
Matches 128; Conservative 27; Mismatches 46; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALIELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML 60
D 1 MARLINESVIDAELNENETGIDGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEVL 60
QY 61 DRHTHFCPLKGSWQDFLRNKAQKFRALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120
D 61 ARHHDYSLPAAGESWQDFLRNKAQKFRALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120
QY 121 COQGSLENAIALYSAVGHFTLCVLEQDEHOVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
D 121 TENGFSRLDGLYSAVGHFTLCVLEQDEHOVAKEREETPTTDSMPPLLRQAIELEFDHQ 180
QY 181 GAEPAPFLGLELIICGLEKQKLCESGS 201
D 181 DGEQAFHLGLESIRGEVQL 201

RESULT 3
TER8_PASMU
ID TER8_PASMU STANDARD; PRT; 207 AA.
AC P51561;
DT 01-OCT-1996 (REL. 34, CREATED)
```

```
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS H.
GN TETR.
OS PASTEURILLA MULTOCIDA.
OG PLASMID P5111.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC PASTEURILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-P2862;
RX MEDLINE; 94153051.
RL HANSEN L.M., MCMURRY L.M., LEVY S.B., HIRSH D.C.;
RL ANTIMICROB. AGENTS CHEMOTHER. 37:2699-2705(1993).
CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETA) PROMOTER OPERATOR SITES.
CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; U00792; G392872; -.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;
KW ANTI-BIOTIC RESISTANCE; PLASMID.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT METAL 100 100 SIMILARITY).
FT SEQUENCE 207 AA; 23156 MW; A51DA28B CRC32;

Query Match 61.4%; Score 663; DB 1; Length 207;
Best Local Similarity 63.2%; Pred. No. 1.7e-55;
Matches 129; Conservative 29; Mismatches 44; Indels 2; Gaps 1;

QY 1 MSRLDKSVINSALIELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAEML 60
D 1 MAKLDKEQVIDDALILLNEVGIEGLTTRNVAQKIGVEQPTLYWHVKNKRALDLALAEIL 60
QY 61 DRHTHFCPLKGSWQDFLRNKAQKFRALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120
D 61 QKHHHVLPNETWQDFLRNKAQKFRALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120
QY 121 COQGSLENAIALYSAVGHFTLCVLEQDEHOVAKEREETPTTDSM--PPLLRQAIELEFD 178
D 121 CDAGFSLSQAVYALSSIAHFTLGSVLETOEQHSQKEREKVEITDVAIPPPLLTQAVAIMD 180
QY 179 HQGAEPAPFLGLELIICGLEKQKLC 202
D 181 SDNGDAAFLEFVLDVYMISGLETVLK 204

RESULT 4
TER4_ECOLI
ID TER4_ECOLI STANDARD; PRT; 217 AA.
AC P09164;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS D.
GN TETR.
OS ESCHERICHIA COLI, AND SALMONELLA ORDONEZ.
OG PLASMID PIP173.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.COLI; STRAIN=J53-1; PLASMID=RA1;
RX MEDLINE; 85037938.
RA UNGER B., KLOCK G., HILLEN W.;
```

RL NUCLEIC ACIDS RES. 12:7693-7703(1984).
RN (2)
RP SEQUENCE FROM N.A.
RC SPECIES-S ORDONEZ; STRAIN-BM2000; PLASMID-PIP173;
RX MEDLINE; 93204906.
RA ALLARD J.D., GIBSON M.L., VU L.H., NGUYEN T.T., BERTRAND K.P.;
RL MOL. GEN. GENET. 237:301-305(1993).
RN (3)
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS), AND PARTIAL SEQUENCE.
RX MEDLINE; 94204640.
RA HINRICHS W., KISKER C., DUEVEL C., MUELLER A., TOVAR K., HILLEN W.,
RA SAENGER W.;
RL SCIENCE 264:418-420(1994).
CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT: ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETR) PROMOTER OPERATOR SITES.
CC -!- SUBUNIT: HOMODIMER.
CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; X01083; G43054; -;
DR EMBL; X55876; G49075; -;
DR PIR; S07359; S07359.
DR PIR; S30287; S30287.
DR PDB; 2TCT; 03-APR-96.
DR PDB; 2TRT; 20-JUN-96.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
RW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;
KW ANTI-BIOTIC RESISTANCE; 3D-STRUCTURE.
FT INIT_MET 0
FT DNA_BIND 25 44 H-T-H MOTIF (POTENTIAL).
FT SITE 63 63 INVOLVED IN BINDING TO [MG-TC]+.
FT METAL 99 99 MAGNESIUM (OF [MG-TC]+ COMPLEX).
SQ SEQUENCE 217 AA; 24287 MW; 830EF8EB CRC32;

Query Match 61.1%; Score 660; DB 1; Length 217;
Best Local Similarity 63.5%; Pred. No. 3.5e-55;
Matches 127; Conservative 27; Mismatches 46; Indels 0; Gaps 0;

QY 2 SRLDKSVINSALELLNEVEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALATEMLD 61
DB 1 ARLNRESVIDAALELLNETGIDGUTTRKLAQKLGIEOPTLYWHVKNKRALLDALAVEILA 60

QY 62 RHHTHFCPLKGSQWDFLRNKAQKSHRGNKAKVHSDTRPTEKQYETLENQALFLC 121
DB 61 RHHDYSLPAGESQWDFLRNKAQKSHRGNKAKVHSDTRPTEKQYETLENQALFLC 120

QY 122 QOGFSLNALVALSAVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQALFDFHQG 181
DB 121 ENGFSRLDGLYSAISAVSHFTLGVLEQOEHTAALTDRAAPDENLPLLRQALQIMDSDD 180

QY 182 AEPALFGLLELIICGLEKOL 201
DB 181 GEQAPLHGLSLIRGFEVOL 200

RESULT 5
TER3_ECOLI STANDARD; PRT; 219 AA.
AC P03039;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS C.
GN TETR.
OS ESCHERICHIA COLI.
OG PLASMID PSC101.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN (1)

RP SEQUENCE FROM N.A.
RX MEDLINE; 88216101.
RA BROW M.A.D., PESIN R., SUTCLIFFE J.G.;
RL MOL. BIOL. EVOL. 2:1-12(1985).
CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT: ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETR) PROMOTER OPERATOR SITES.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; M36272; G150946; -;
DR PIR; A03575; RPECYS.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
RW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;
KW ANTI-BIOTIC RESISTANCE.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT SIMILARITY).
FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
FT SIMILARITY).
SQ SEQUENCE 219 AA; 24174 MW; 6C9E1828 CRC32;

Query Match 47.4%; Score 512; DB 1; Length 219;
Best Local Similarity 51.7%; Pred. No. 3e-41;
Matches 105; Conservative 33; Mismatches 61; Indels 4; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALATEML 60
DB 1 MNKLOREAVIRTALELLNDVGMGLTTRLAERLGVQOPALYWHFKNKRALLDALAEAML 60

QY 61 DRHHTHFCPLKGSQWDFLRNKAQKSHRGNKAKVHSDTRPTEKQYETLENQALFLC 120
DB 61 TINHTHPTDRDDWRSEFLKGNACSFRRALLAYRDGARIHAGTRPAAPQMEKAQURFL 120

QY 121 COOGFSLNALVALSAVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQALFEL 176
DB 121 CDAGFSAGATYALMAISYFTVGAVLEQOASEADAERGEDQLTTSASTMPARLQSAWKI 180

QY 177 FDHQGAEPALFGLLELIICGLEK 199
DB 181 VYEGGPDAAFERGLALIIGGLEK 203

RESULT 6
TER1_ECOLI STANDARD; PRT; 216 AA.
AC P03038;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS A (TRANSPONSON 1721).
GN TETR.
OS ESCHERICHIA COLI.
OG PLASMID RPI, AND PLASMID RP4.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN (1)
RP SEQUENCE FROM N.A.
RC TRANSPONSON-TN1721;
RX MEDLINE; 92192465.
RA ALLMEIER H., CRESNAR B., GRECK M., SCHMITT R.;
RL GENE 111:11-20(1992).
RN (2)
RP SEQUENCE FROM N.A.
RC PLASMID-PLASMID RP4;
RA TRUENMAN P., SHARPE G.S., BARTH P.T.;
RL SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN (3)
RP SEQUENCE FROM N.A.
RC TRANSPONSON-TN1721; PLASMID-RP1;
RX MEDLINE; 83299270.

	Query Match	43.8%	Score 473;	DB 1;	Length 210;	
	Best Local Similarity	47.3%	Pred. No. 1.3e-37;			
	Matches	95;	Conservative	32;	Mismatches 72;	Gaps 1;
Qy	1 MSRLDKSVINSALELLNEVGIEGLTTRKLAQKLGVQEPTLYWHVKNKRALLDALAEML	60				
Ddb	1 : : :: :: :					
	1 MTKLDGTVIAAGLELLNEVGMDSLTTRKLABRLKVQPALYWHFQNKRALLDALPEAML	60				
Qy	61 DRHHTFCPLKGESWODFLNKKAKSCFALLSHRNGAKVHSDTREKQYETLENQLAEFL	120				
Ddb	11 : : : : : : : : : : : : : :					
Qy	61 RERHTRSLPENEDRWRFLENALSFTALLSTROGARIHAGTREFNFGTATQTIRFL	120				
Ddb	1 : : : : : : : : : : : : : :					
Qy	121 CQQGFSLNALYALSAYGHFTTLGCVLDEQHGVAKEREETPTDTSMP--LLRQAIELFD	178				
Ddb	1 : : : : : : : : : : : : : :					
Qy	121 CAEGFCPKRAVMALRAYSHVVGVSGVLEQQASDADERVPDRPDVSEQAPSFHLVLFHELE	180				
Ddb	179 HOGAEPAPFLGLELICGLEK	199				
Qy	181 TDGMDAAFNFLDSLIAGFER	201				
Ddb						
	RESULT 9					
	Y07H_MYCTU					
ID	Y07H_MYCTU	STANDARD;	PRT;	261 AA.		
AC	Q11023;					
DT	01-OCT-1996 (REL. 34, CREATED)					
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)					
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)					
DE	HYPOTHETICAL TRANSCRIPTIONAL REGULATOR CY02B10.17C.					
GN	MYCY02B10.17C.					
OS	MYCOBACTERIUM TUBERCULOSIS.					
OC	PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-H37RV;					
RA	MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A.; WALSH S.V.;					
RL	SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.					
CC	-!- SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL REGULATORS.					
CC	EMBL; Z75555; -.					
DR	PROSITE; PS01081; HTH_TETR_FAMILY; FALSE_NEG.					
DR	KW HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING.					
FW	DNA_BIND 38 57 H-T-H MOTIF (POTENTIAL).					
FT	SEQUENCE 261 AA; 28253 MW; 8CBBC98 CRC32;					
SQ						
	Query Match	15.6%	Score 168.5;	DB 1;	Length 261;	
	Best Local Similarity	27.0%;	Pred. No. 7.8e-09;			
	Matches	61;	Conservative	32;	Mismatches 78;	Gaps 6;
Qy	4 LDKSKVINSALELLNEVGIEGLTTRKLAQKLGVQEPTLYWHVKNKRALLDALAEMLDRH	63				
Db	16 INPEDIIAGFAELAAQOVSDINLSMPLLGKHGLGVGTSIYWYFRKDDLLNAMTDRALSKI	75				
Qy	64 HTHFCPLKGESWODFLNKKAKSR-----CALL-----SHRNGAKVHSDTRP	105				
Ddb	76 VFATPIEAGDWREILRNHARSWKTFADNPVLCDLILRAALSPKTAFLGA-----	128				
Qy	106 TEKQYETLENQLAFLCQQCGFSLENALYALSAVGHFTTCGVLEDO-----E	150				
Db	128 -----GEMEKAIANIVTAGLSLEDAFIYSVAHSVVRGSVVLDRLSRKSQSAGSGPSAIE	182				
Qy	151 HOVAKEERTPTDTSMPPLLRAQIELFDHOGA--EPAFLFGLELI	194				
Ddb	183 HPVAIDPAT-----PLAHATGRGHRIAGDPETNEFYGLECIL	221				
	RESULT 10					
	MTRR_NEIGO					
ID	MTRR_NEIGO	STANDARD;	PRT.	210 AA.		
AC	B36827.					

FT DNA_BIND 67 86 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 236 AA; 26204 MW; AD1F64FC CRC32;

Query Match 8.6%; Score 92.5; DB 1; Length 236;
Best Local Similarity 25.2%; Pred. No. 0.098;
Matches 35; Conservative 18; Mismatches 63; Indels 23; Gaps 5;

QY 5 DSKVINSALLELNEVGIEGLTRKLAQKLVGEQPTLYHVHVKNRALLDALAIEMLDRL-- 63
DB 46 DRDAALDKAMKLFVHGYEATSLADLVATGCAKAPTYAEFTNKGLFRAVLDRYIDREA 105
QY 63 --HHTH-FCPLKG--ESWQDFLRNKAISF-----RCALLSHENGAKVHSDR 104
DB 106 AKHEAQLFCCKESVESALADYFAAIANCFTSKDTPAGCFMINNCTTSPDSDGIANTLKS 165
QY 105 PTEKQYETLENQALFLCOQ 123
DB 166 RHAMQERTLQO---FLCOR 181

RESULT 12
UIIDR_ECOLI STANDARD; PRT; 196 AA.
AC Q39431;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE UID OPERON REPRESSOR (GUS OPERON REPRESSOR).
GN UIDR OR GUSR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
CC [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA WILSON K.J.; JEFFERSON R.A.;
RL SUBMITTED (OCT-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA BLATTNER F.R.; PLUNKETT G. III; MAYHEW G.F.; PERNA N.T.; GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA AIBA H.; BABA T.; FUJITA K.; HAYASHI K.; HONJO A.; HORIIUCHI T.;
RA IKEMOTO K.; INADA T.; ISONO K.; ISONO S.; ITOH T.; KANAI K.; KASAI H.;
RA KASHIMOTO K.; KIM S.; KIMURA S.; KITAGAWA M.; KITAKAWA M.; MAKINO K.;
RA MASUDA S.; MIKI T.; MIZOBUCHI K.; MORI H.; MOTOMURA K.; NAKAMURA Y.;
RA NASHIMOTO H.; NISHIO Y.; OSHIMA T.; SAITO N.; SAMEI G.; SEKI Y.;
RA TAGAMI H.; TAKEMOTO K.; WADA C.; YAMAMOTO Y.; YANO M.;
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: REPRESSOR FOR THE UIDRABC (GUSRABC) OPERON.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; M14641; G868019; -.
DR EMBL; AE000257; G1787904; -.
DR EMBL; D90805; G1742672; -.
DR EMBL; EG12667; UIDR.
DR PROSITE; PS01081; HTH TETR FAMILY; 1.
KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.
FT DNA_BIND 33 52 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 196 AA; 21799 MW; 6C1840FC CRC32;

Query Match 8.5%; Score 91.5; DB 1; Length 196;
Best Local Similarity 24.0%; Pred. No. 0.097;
Matches 18; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

QY 6 KSKVINSALLELNEVGIEGLTRKLAQKLVGEQPTLYHVHVKNRALLDALAIEMLDRLHHT 65
DB 13 RTRILNAAREIFSENGFHSASMAKICKSCAISPGLYHHFISKEALIQAILQDOERALA 72

QY 66 HFC-PLKGSWQDFL 79
DB 73 RREPPIEGIHFDYM 87

RESULT 13
PHSL_IPOBA STANDARD; PRT; 955 AA.
ID PHSL_IPOBA
AC P27598;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ALPHA-GLUCAN PHOSPHORYLASE, L ISOZYME PRECURSOR (EC 2.4.1.1)
DE (STARCH PHOSPHORYLASE L).
OS IPOMOEA BATATAS (SWEET POTATO) (BATATE).
OC EUKARYOTA; PLANTA; EMERYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC SOLANALES; CONVULVACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RA LIN C.T.; YEH K.W.; LEE P.D.; SU J.C.;
RL PLANT PHYSIOL. 0:0-0(1992).
CC -!- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN
CC CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
CC THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
CC PROPERTIES.
CC -!- CATALYTIC ACTIVITY: (1,4-ALPHA-D-GLUCOSYL)(N) + ORTHOPHOSPHATE =
CC (1,4-ALPHA-D-GLUCOSYL)(N-1) + ALPHA-D-GLUCOSE 1-PHOSPHATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST; AMYLOPLAST.
CC -!- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.
DR EMBL; M64362; G168276; -.
DR HSP; P00489; 9GPB.
DR PROSITE; PS0102; PHOSPHORYLASE; 1.
KW TRANSFERASE; GLYCOSYLTRANSFERASE; CARBOHYDRATE METABOLISM;
KW ALLOSTERIC ENZYME; PYRIDOXAL PHOSPHATE; TRANSIT PEPTIDE; CHLOROPLAST;
KW AMYLOPLAST; MULTIGENE FAMILY.
FT TRANSIT 1 43 CHLOROPLAST (POTENTIAL).
FT CHAIN 44 955 ALPHA-GLUCAN PHOSPHORYLASE, L ISOZYME.
FT BINDING 801 801 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 955 AA; 108520 MW; 86E3D50F CRC32;

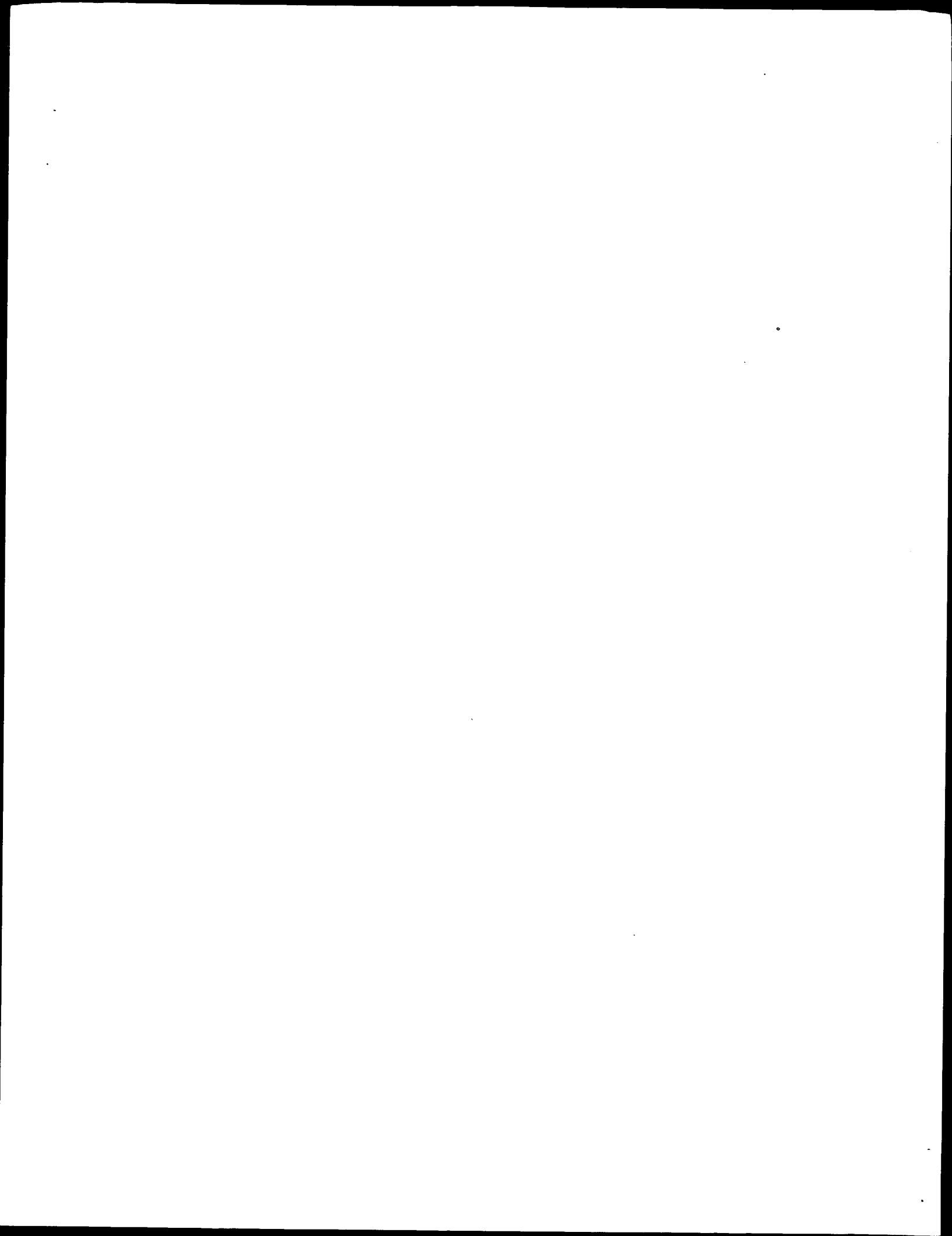
Query Match 8.0%; Score 86; DB 1; Length 955;
Best Local Similarity 30.1%; Pred. No. 2.2;
Matches 56; Conservative 19; Mismatches 67; Indels 44; Gaps 10;

QY 1 MSRL-----DKSVINSALLELNEVGIEGL--TTRKLAQKLV-EOPTLYHVH---- 46
DB 1 MSRLSGITPRADRSQFONPRLEIAVDPRTAGLQRTKTLVLCVLDKTKQTIQHVYTE 60
QY 46 KNRALLDALAIEMLDRLHHTFCPLKGSWQDFLRNKAISF-----AKSFRCALLSHENGAKVH 100
DB 61 KNEGTLLDAASIASSIKYHAEFSP--AFSPERFELPKAFATQSVRDALIVNNA---- 115
QY 101 SDTRPTEKQYETLENQALFLCO-----QGFSLNAYALVAVGHFT-----LGCVLDEQEH 151
DB 115 -----TYDYEXLNMKQAYLSMEFLQGRALLNAGLNLGTGEYAEALNKLGHNLN---- 167
QY 152 QVAKEE 157
DB 167 -VASKE 171

RESULT 14
THTL_SCHPO STANDARD; PRT; 577 AA.
ID THTL_SCHPO
AC Q09684;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NUCLEAR FUSION PROTEIN THTL.

RESULT	15	
KINH_HUMAN		
ID	KINH_HUMAN	STANDARD; PRT; 963 AA.
AC	P3176:	
DT	01-OCT-1993	(REL. 27, CREATED)
DT	01-OCT-1993	(REL. 27, LAST SEQUENCE UPDATE)
DT	15-JUL-1998	(REL. 36, LAST ANNOTATION UPDATE)
DE	DE KINESIN HEAVY CHAIN (UBIQUITOUS KINESIN HEAVY CHAIN) (UKHC).	
OS	KNS1 OR KNS.	
OS	OS HOMO SAPIENS (HUMAN).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC	EUTHERIA; PRIMATES.	
RC	[1]	
RC	SEQUENCE FROM N.A.	
RC	TISSUE=PLACENTA;	
RC	MEDLINE; 92299683.	
EX	NAVONE F., NICLAS J., HOM-BOOHER N., SPARKS L., BERNSTEIN H.D.,	
RA	RA MCCAFFREY G., VALE R.D.;	
RA	J. CELL BIOL. 117:1263-1275(1992).	
RL	[2]	
RL	X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-349.	
RP	MEDLINE; 96193066.	
RA	KULL F.J., SABLIN E.P., LAU R., FLETTERICK R.J., VALE R.D.;	
RL	RL NATURE 380:550-555(1996).	
RL	[3]	
RP	TISSUE SPECIFICITY.	
RP	MEDLINE; 94242426.	
RA	RA NICLAS J., NAVONE F., HOM-BOOHER N., VALE R.D.;	
RA	RA NEURON 12:1059-1072(1994).	
CC	-1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING	

Search completed: June 9, 1999, 13:01:03
Job time: 9122 sec



GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 13:00:33 ; Search time 28.57 Seconds
(without alignments)
399,721 Million cell updates/sec

Title: US-08-486-814-19
Perfect score: 1080
Sequence: 1 MSRLDKSVINSALELLNEV.....FGLELIICGLEKQLKCESGS 207

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database : SPREMBL_8.*
1: sp_fungi.*
2: sp_human.*
3: sp_invertebrate.*
4: sp_mammal.*
5: sp_mhc.*
6: sp_organelle.*
7: sp_phage.*
8: sp_plant.*
9: sp_bacteria.*
10: sp_rodent.*
11: sp_virus.*
12: sp_vertebrate.*
13: sp_unclassified.*
14: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	597	55.3	126	9	Q59466	Q59466 haemophilus
2	147.5	13.7	259	9	Q53901	Q53901 streptomyces
3	124.5	11.5	242	9	O52558	O52558 amycolatops
4	123	11.4	191	9	O34892	O34892 bacillus su
5	105.5	9.8	210	9	O59608	O59608 neisseria g
6	103.5	9.6	192	9	O67927	O67927 aquifex aeo
7	102	9.4	189	9	O34843	O34843 bacillus su
8	99	9.2	196	9	O53789	O53789 mycobacteri
9	98.5	9.1	186	9	O51597	O51597 pseudomonas
10	96.5	8.9	246	9	O53295	O53295 mycobacteri
11	94	8.7	179	9	O67157	O67157 aquifex aeo
12	93.5	8.7	236	9	O73186	O73186 streptococ
13	93	8.6	217	9	O87854	O87854 streptomyces
14	91.5	8.5	213	9	O53310	O53310 mycobacteri
15	91.5	8.5	194	9	O07001	O07001 bacillus su
16	89	8.2	202	9	O51730	O51730 pseudomonas
17	88.5	8.2	2472	10	P70477	P70477 rattus norv
18	88.5	8.2	2472	10	O88663	O88663 rattus norv
19	86.5	8.0	2477	2	Q13186	Q13186 homo sapien
20	86.5	8.0	435	2	O75901	O75901 homo sapien
21	86.5	8.0	185	9	O70020	O70020 staphylococ
22	86	8.0	340	8	O42863	O42863 ipomoea bat
23	86	8.0	488	9	O54830	O54830 streptococ
24	86	8.0	581	9	O54835	O54835 streptococ
25	86	8.0	205	9	O51516	O51516 pseudomonas
26	84	7.8	539	9	O54719	O54719 streptococ
27	83.5	7.7	216	9	P95100	P95100 mycobacteri
28	83	7.7	2748	1	O03767	O03767 saccharomyc
29	83	7.7	198	9	P96676	P96676 bacillus su

Q59802 staphylococ
Q29994 archaeoglob
O18399 drosophila
Q91290 pleurodeles
O34970 bacillus su
O19126 macaca fasc
O19127 macaca mula
O25951 helicobacte
O86852 streptomyce
O53165 mycobacteri
O07388 mycobacteri
O90631 gallus gall
O59306 clostridium
P94548 bacillus su
O15087 homo sapien
P79138 cercopithec

ALIGNMENTS

RESULT 1

Q59466
ID O59466 PRELIMINARY; PRT; 126 AA.
AC O59466;
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE 30S RIBOSOMAL PROTEIN S21.
OS HAEMOPHILUS PARAINFLUENZA.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87137315.
RA HEUER C., HICKMAN R.K., CURIALE M.S., HILLEN W., LEVY S.B.;
RT "Constitutive expression of tetracycline resistance mediated by a
RT Tn10-like element in Haemophilus parainfluenzae results from a
RT mutation in the repressor gene."
RL J. BACTERIOL. 169:990-994(1987).
CC -!- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
DR EMBL; M15539; G148989;
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
DR PFAM; PF00440; tetr; 1.
KW RIBOSOMAL PROTEIN.
SQ SEQUENCE 126 AA; 14541 MW; 107F233B CRC32;

Query Match 55.3%; Score 597; DB 9; Length 126;
Best Local Similarity 92.0%; Pred. No. 1.3e-45;
Matches 115; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYHWYKVRALLDALAIEML 60
|||||
Db 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYHWYKVRALLDALAIEML 60

QY 61 DRHHTHFCPLKESQWDFLRNKAQKFRALLSHRNKAVHSDTRTEKQYETLENQALFL 120
|||||
Db 61 DRHHTHFCPLKESQWDFLRNKAQKFRALLSHRNKAVHSDTRTEKQYETLENQALFL 120

QY 121 COQGF 125
:|

Db 121 ANKVF 125

RESULT 2

Q53901 PRELIMINARY; PRT; 259 AA.
ID Q53901
AC Q53901;
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE ORF1-4 (ACTII).

```

GN ACTII.
OS STREPTOMYCES COELICOLOR.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; STREPTOMYCETES;
OC STREPTOMYCETACEAE; STREPTOMYCES.
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE: 91347376.
RA FERNANDEZ-MORENO M.A., CABALLERO J.L., HOPWOOD D.A., MALPARTIDA F.;
RT "the act cluster contains regulatory and antibiotic export genes,
RT direct targets for translational control by the bldA trna gene of
RT Streptomyces.";
RL CELL 66:769-780(1991).
DR EMBL: M64683: G453360: -.
DR PFAM: PF00440: tetr; 1.
SQ SEQUENCE 259 AA; 28268 MW; 16ECA071 CRC32;

Query Match 13.7%; Score 147.5; DB 9; Length 259;
Best Local Similarity 26.0%; Pred. No. 1e-05;
Matches 57; Conservative 35; Mismatches 90; Indels 37; Gaps 8;

QY 4 LDKSKVINSALELLNVGVIEGTTTRKLAQKLGVSQPTLYWHVKNKRALLDAL-----AI 57
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
31 LTODRIVVTALGTLDAEGLDALSMLRLAQELKTHASLYAHVGNRDELLDVFIVLTV 90
QY 58 EMLDRHHTHFCPLKGSWODFNRNKAQKSFRCALLSHRNKAKVHSDTRPT-EKQYETLE 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
91 EVPE-----PEPGR-WAEQVKEMCRSLRMFLAHRDLARIAIDRVPLGPNMGWGMERT 142
QY 117 LAFLCQOGFSLNAYLSAVGHFTLGCVLDDQEHQVAKEE-----RET 160
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
143 MNLLRSGGLHDELAAYGGDLLSTFTVTAELQSSRNPTGEGQAGYFADQLHGLYKSL 202
QY 161 PTTDSMPPLLRQA--IELFDHOGAEPALFGLLELIIGL 197
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
203 PAT-SFPLVHLGAPITSLD---SDRRELGLEIIAGL 237

RESULT 3
O52558 052558 PRELIMINARY; PRT; 242 AA.
ID AC O52558;
AC O52558;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DI 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE REPRESSOR.
GN RIFQ.
OS AMYCOTATOPSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; PSEUDONOCARDIACEAE;
OC AMYCOTATOPSIS.
[1]
RN RP SEQUENCE FROM N.A.
RX STRAIN-S699;
RA KIM C.G., YU T.W., FRYHLE C., HANDA S., FLOSS H.G.;
RL J. BIOL. CHEM. 0:0-0(1998).
[2]
RN RP SEQUENCE FROM N.A.
RX STRAIN-S699;
RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., YU T.W.;
RA TAYLOR M., HOFFMANN D., KIM C.G., ZHANG X., HUTCHINSON C.R.,
RA FLOSS H.G.;
RL CHEM. BIOL. 5:0-0(0002).
[3]
RN RP SEQUENCE FROM N.A.
RX STRAIN-S699;
RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., HUTCHINSON C.R.,
RA FLOSS H.G.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF040570: G2792330: -.
SQ SEQUENCE 242 AA; 27049 MW; 324EB570 CRC32;

Query Match 11.5%; Score 124.5; DB 9; Length 242;

```

QY		179 HQGAPPAFLF 188	:	
Dd		156 DLDKETAVIF 165	:	
RESULT	6			
O67927		PRELIMINARY;	PRT;	192 AA.
ID	O67927			
AC	O67927			
DT	01-AUG-1998 (TREMBREL. 07, CREATED)			
DT	01-AUG-1998 (TREMBREL. 07, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)			
DE	TRANSCRIPTIONAL REGULATOR (TETR/ACRR FAMILY).			
GN	ACR2.			
OS	AQUIFEX AEOLICUS.			
OC	BACTERIA; AQUIFICAE; AQUIFICACEAE; AQUIFEX.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=VF5;			
RA	MEDLINE; 98196666.			
RA	DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,			
RA	GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,			
RA	FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;			
RT	"The complete genome of the hyperthermophilic bacterium Aquifex			
RL	aerolicus.";			
RL	NATURE 392:353-358(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=VF5;			
RA	DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,			
RA	GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,			
RA	FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;			
RL	SUBMITTED (JUL-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.			
DR	ENBL; AE000776; G2984362;			
DR	PROSITE; PS01081; HTH.TETR.FAMILY: 1.			
SQ	SEQUENCE 192 AA; 22259 MW; D2E83A94 CRC32;			
Query Match	9.6%; Score 103.5; DB 9; Length 192;			
Best Local Similarity	22.7%; Pred. No. 0.052;			
Matches	45; Conservative 35; Mismatches 67; Indels 51; Gaps			
Qy	6 KSKVINSALELLNEVGIEGTTRKLAQGLGVQPTLYWHVNKK---RALDLALAEMLD 61 : :			
Dd	15 KEKTLSSALKLFSGKGFKETTKIDIAKEVGITEGAIRHFTSKSEETIKLSLESITKEL-- 73 : :			
Qy	62 RHHTHFCLKGESWQDFLR-----NKAKSFRCALLSHRNKAKVHSDRTPTEKOY 110 : : : : : : : : : : : : : : : : : : : :			
Dd	73 RHKLEVALQRGETDEETILESIVTDLIYAFSNPESFRFLNYH-----LLKEY 120 : : : : : : : : : : : : : : : : : : : :			
Qy	111 ETLENQ-----LAFPCOQGFSLENALY---ALSAGVHFTLCGVLEDQEQHVAKKER--- 159 : : : : : : : : : : : : : : : : : : : :			
Dd	121 GEVKNLPGCELILKFL-----NGLYKRKLKTYPEIALAVVTGSVERVFIKERNFLD 172 : : : : : : : : : : : : : : : : : : : :			
Qy	159 --ETPTDSNPPLLRQA 174 : : : : : : : : : : : : : : : : : : : :			
Dd	173 YDEETIKELKKVLKSAI 190 : : : : : : : : : : : : : : : : : : : :			
RESULT	7			
O34643		PRELIMINARY;	PRT;	189 AA.
ID	O34643			
AC	O34643;			
DT	01-JAN-1998 (TREMBREL. 05, CREATED)			
DT	01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)			
DE	YVXB.			
GN	YVXB.			
OS	BACILLUS SUBTILIS.			
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;			
OC	BACILLUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			

```

OC MYCOBACTERIUM.
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RC RAMLIN N., CHURCHER C.M.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RL [2]
RN SEQUENCE FROM N.A.
RN STRAIN-H37RV;
RC COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RL [3]
RN SEQUENCE FROM N.A.
RN STRAIN-H37RV;
RC MEDLINE; 96181548.
RC PHILIPP W.J., POULST S., EIGMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37Rv, and comparison with Mycobacterium
RT leprae".
RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMB; AL021943; E1253107; -.
SQ SEQUENCE 196 AA; 21262 MW; 17D86001 CRC32;

Query Match          9.2%; Score 99; DB 9; Length 196;
Best Local Similarity 25.8%; Pred. NO. 0.13;
Matches 40; Conservative 27; Mismatches 72; Indels 16; Gaps

QY 2 SRLDKSVINSALELLNEVGIEGLTRTKLAQKIQEOPTLYWHVKNKRALLDALATEMLD 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 AKLSRESIVEGALTFLDREGWDSLTINALQTGKGPLSYNHVDSLEDLRRAVRVID 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 RHHTHFCL-KGSWQDFLNKAKSFRCALMSHRNKAHVSDTR----PTEKQYETLEN- 116
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 DIITMLNVGAGRARDVAVLVWAGYRS--YAHHPGRYSFTMPJGGDDPEYTAATRG 122
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

01-NOV-1996 (CREATED)
01-NOV-1996 (TREMREL, 01, LAST SEQUENCE UPDATE)
01-NOV-1998 (TREMREL, 08, LAST ANNOTATION UPDATE)
CAM REPRESSOR.
CAMR.
PSEUDOMONAS PUTIDA.
OS
PLASMID CAM.
OG
BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;
OC
PSEUDOMONAS.
OC
[1]
SEQUENCE FROM N.A.
RRP
STRAIN=PPGI;
RCR
MEDLINE; 86223770.
KOGA H.; ARAMAKI H., YAMAGUCHI E., TAKEUCHI K., HORIUCHI T.,
GUNSALUS I.C.;
"camr, a negative regulator locus of the cytochrome P-450cam
hydroxylase operon";
J. BACTERIOL. 166:1089-1095(1986).
[2]
SEQUENCE FROM N.A.
RRP
STRAIN=PPGI;
RCR
MEDLINE; 93326643.
ARAMAKI H., KOGA H., SAGARA Y., HOSOI M., HORIUCHI T.;
"Complete nucleotide sequence of the 5-exo-hydroxycamphor
dehydrogenase gene on the CAM plasmid of Pseudomonas putida (ATCC
17453).";
J. BACTERIOL. 174:531-536(1990).

GUNSALUS I.C.;
"camk, a negative regulator locus of the cytochrome P-450cam
hydroxylase operon.";
J. BACTERIOL. 166:1089-1095(1986).
[2]
SEQUENCE FROM N.A.
STRAIN=PPG1:
MEDLINE; 93326643.
ARAKASHI H., KOGA H., SAGARA Y., HOSOI M., HORIUCHI T.;
"Complete nucleotide sequence of the 5'-exo-hydroxycamphor
dehydrogenase gene on the CAM plasmid of *Pseudomonas putida* (ATCC
17453)." ;


```

RESULT 13
CG7834
ID O87854 PRELIMINARY; PRT: 217 AA.
AC O87854;
DT 01-NOV-1998 (TREMBLREL, 08, CREATED)
DT 01-NOV-1998 (TREMBLREL, 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)
DE PUTATIVE TRANSCRIPTIONAL REGULATOR.
GN SCBA6_20C.
OS STREPTOMYCES COELICOLOR.
OC BACTERIA, FIRMICUTES; ACTINOBACTERIA; STREPTOMYCETES;
OC STREPTOMYCETACEAE; STREPTOMYCETES.
RN [1]
RN RP SEQUENCE FROM N.A.
RN RC STRAIN-A3(2);
RA SEEGER K.J., HARRIS D.;
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]

```

30 BACTILLUS SUBTILIS.

BACILLUS SUBTILIS.

Fri Jun 11 11:58:26 1999

Search completed: June 9, 1999, 13:00:33
Job time: 9152 sec

